MULTILEVEL MODELLING NEWSLETTER

The Multilevel Models Project
Mathematical Sciences
Institute of Education, University of London
20 Bedford Way, London WC1H 0AL, ENGLAND
E-mail: m.yang@ioe.ac.uk
Web site http://www.ioe.ac.uk/multilevel/
Tel: +44 (0)171 612 6682 Fax: +44 (0)171 612 6686

Vol. 10 No. 1 **February, 1998**

Advanced Training Workshops in multilevel modelling in 1998 :

Under phase II of the Analysis of Large and Complex Datasets (ALCD) programme of the Economic and Social Research Council (ESRC) in the UK, five advanced workshops for established researchers in the fields of Education, Public Health, Geography/Environment, Political Science and Demography will be held. They are two-day workshops by invitation at the Institute of Education in London. The workshop dates and organisers are

- 2-3 Feb. by *Harvey Goldstein* (Education);
- 18-19 March by *Ian Langford* (Environment);
- 7-8 May by *Alastair Leyland* (Public Health);
- 15-16 October by *Anthony Heath* (Political);
- 3-4 Dec. by *Ian Diamond* (Demography).

More forthcoming Events:

• **6 April**, a one-day workshop on Markov Chain Monte Carlo estimation using *MLwiN* will take place at Institute of Education, University of London. *David Draper* and *William Browne* (B.Browne@maths.bath.ac.uk) from Bath University will be organising it.

• **May 21-22**, Statistics Canada Symposium 1998 will be held in Ottawa, on the tipic of longitudinal analysis for complex surveys. Multilevel modelling techniques and applications to longitudinal survey data is one of the topics (symposium@statcan.ca).

• **Workshop in Toronto**, a two-day workshop using *MLwiN* will take in mid May in Toronto. Enquiries to Lorna Earl (learl@oise.utoronto.ca).

• **October 20**, a one-day meeting in London will be held on Applications of Random Effects / Multilevel Models to Categorical Data in Social Science and Medicine. This meeting is sponsored jointly by Statistics in Society (JRSS (A)), the Social Statistics and Medical Sections of the Royal Statistical Society, and the ALCD programme of the ESRC. The aims of the meeting are to bring together statisticians in the UK and abroad working in these areas, to show improvement of statistical inference by means of applying these modern methods, to demonstrate to interested non-statisticians the value of these approaches and to produce a set of papers suitable for publication in JRSS (A). The organizing committee consists of *Ian Plewis*, *Gillian Raab*, *Fred Smith*, *Patrick Heady* and *John Wakefield* (i.plewis@ioe.ac.uk).

Also In This Issue

A conditioned iterative generalised least squares estimator (CIGLS)

Evaluating effectiveness of universities through the analysis of student careers

Publications on multilevel analysis in 1997

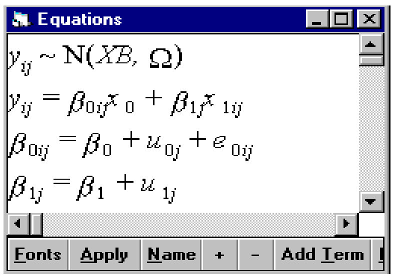
Introducing *MLwiN* -- The Windows based successor to *MLn*

E·S·R·C ECONOMIC

Some features of *MLwiN*: a visual interface for multilevel modelling

The equations window

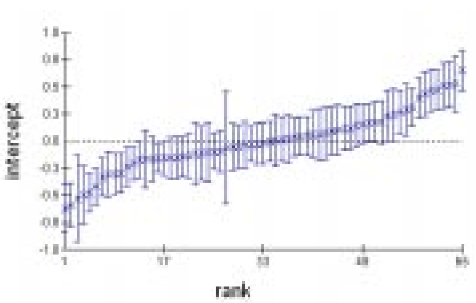
In *MLwiN* you can specify models in three different ways. The original command interface of MLn can be used - and in release 1.0 you will have to use this for certain advanced features. A system of dialogue boxes can be completed or you can directly manipulate the elements of the model equation. Consistency is maintained so that, for example, completing a dialogue box also updates the equations.



The *MLwiN* window on the left shows a 2 level random coefficient model with two explanatory variables, a constant, x_0, with no subscripts and an explanatory variable with a level 1 and a level 2 subscript indicating that it varies across each type of unit. *MLwiN* automatically decides which subscripts to use when you define a variable. Note also the Normal distribution assumption - this can be changed for generalised linear models. along with a selection of link functions. The last two lines define the random variables at each level.

Graphing in *MLwiN*

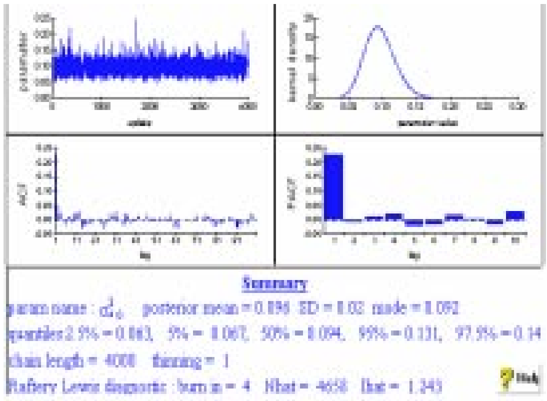
Extensive plotting facilities are available. Any graph can be altered in terms of colour, symbols, lines etc.



We can superimpose graphs, lay them out in patterns (such as trellis plots), label them, identify points or lines on them in terms of data units and copy and paste them to other applications. Several special kinds of graphs are created directly by the software.

For example the 'caterpillar' graph on the left is created from the 'residuals window' and represents a set of ordered 'shrunken' residual estimates of 'school effects' with 95% confidence intervals from a variance components model.

Monitoring MCMC estimation



MCMC methods allow Bayesian models to be fitted with prior parameter distributions By default *MLwiN* sets diffuse priors. Both Gibbs sampling and the Metropolis-Hastings algorithm are used.

We can obtain summary measures and diagnostics by clicking any of these graphs to obtain a window such as the one on the left which shows a kernel density plot, autocorrelation functions and estimates of required chain length etc. for a level 2 variance parameter. Had an informative prior been specified its distribution would have been superimposed on that of the posterior kernel density.

MLwiN has been created by the Multilevel Models Project team based at the Institute of Education, University of London together with various colleagues in other centres, and with support from the ESRC (UK). It is to be launched mid February 1998. It has a Web site (http://www.ioe.ac.uk/mlwin/) where you can browse features, keep up to date with the latest upgrades and releases and download an order form. For further details, send email to mln.order@ioe.ac.uk or phone +44 (0)171 612 6027 or fax +44 (0)171 612 6032. For technical support, send email to m.yang@ioe.ac.uk.

The next issue of the Newsletter will contain a review of *MLwiN*.

Some Publications in 1997 Using Multilevel Models

Adams, R. J., Wilson, M. and Wu, M. (1997). Multilevel item response models: An approach to errors in variables regression. *Journal of Educational and behavioural Statistics* **22**: 47-76.

Agresti, A. (1997). A model for repeated measurements of a multivariate binaryu response. *Journal of the American Statistical Association* **92**: 315-21.

Altman, D. G. and Bland, J. M. (1997). Units of analysis. *British Medical Journal* **314**: 1874.

Bernardinalli, L., Pascutto, C., Best, N. G. and Gilks, W. R. (1997). Disease mapping with errors in covariates. *Statistics in Medicine* **16**: 741-52.

Bigerstaff, B. J. and Tweedie, R. L. (1997). Incorporating variability in estimates of heterogeneity in the random effects model in meta analysis. *Statistics in Medicine* **16**: 753-68.

Candy, S. G. (1997). Estimation in forest yield models using composite link functions with random effects. *Biometrics* **53**: 146-60.

Catalano, P. J. (1997). Bivariate modelling of clustered continuous and ordered categorical outcomes. *Statistics in Medicine* **16**: 883-900.

Chan, J. S. K. and Kuk, A. Y. C. (1997). Maximum likelihood estimation for probit-linear mixed models with correlated random effects. *Biometrics* **53**: 86-97.

Christiansen, C. L. and Morris, C. (1997). Hierarchical Poisson regression modelling. *Journal of the American Statistical Association* **92**: 618-632.

Ecochard, R. (1997). *Random effect models in the statistical analysis of human fecundability data*. PhD theses, University of Cambridge.

Goldstein, H. and Sammons, P. (1997). The influence of secondary and junior schools on sixteen year examination performance: a cross-classified multilevel analysis. *School effectiveness and school improvement.* **8**: 219-230.

Greenland, S. (1997). Second stage least squares versus penalized quasi-likelihood for fitting hierarchical models in epidemiologic analyses. *Statistics in Medicine* **16**: 515-26.

Heitjan, D. F. and Sharma, D. (1997). Modelling repeated-series longitudinal data. *Statistics in Medicine* **16**: 347-355.

Hill, P. W. and Goldstein, H. (1997). Multilevel modelling of educational data with cross classification and missing identification of units. *Journal of Educational and Behavioural statistics (to appear)*.

Hogan, J. W. and Laird, N. M. (1997). Mixture models for the joint distribution of repeated measures and event times. *Statistics in Medicine* **16**: 239-258.

Hogan, J. W. and laird, N. M. (1997). Model based approaches to analysing incomplete longitudinal and failure time data. *Statistics in Medicine* **16**: 259-272.

Larose, D. T. and Dey, D. K. (1997). Grouped random effects models for Bayesian meta analysis. *Statistics in Medicine* **16**: 1817-30.

Lin, X. (1997). Variance component testing in generalised linear models with random effects. *Bimetrika* **84**: 309-25.

McCulloch, C. E. (1997). Maximum likelihoodalgorithms for generalised linear mixed models. *Journal of the American Statistical Association* **92**: 162-70.

Mehrotra, D. V. (1997). Non-iterative robust estimators of variance components in within-subject designs. *Statistics in Medicine* **16**: 1465-80.

Mukhopadhyay, S. and Gelfand, A. E. (1997). Dirichilet process mixed generalised linear models. *Journal of the American Statistical Association* **92**: 633-39.

Pan, H. and Goldstein, H. (1997). Multilevel models for longitudinal growth norms. *Statistics in Medicine* **16**: (To appear).

Pfeffermann, D., Skinner, C. J., Holmes, D., Goldstein, H., et al. (1997). Weighting for unequal selection propbabilities in multilevel models. *Journal of the Royal Statistical Society, B*. (to appear).

Raudenbush, S. W. (1997). Hierarchical models, Bayes, maximum likelihood, multilevel data. *Encycopedia of the social sciences*. New York, Wiley.

Richardson, A. (1997). Bounded influence estimation in the mixed linear model. *Journal of the American Statistical Association* **92**: 154-61.

Sastry, N. (1997). A nested frailty model for survivcal data, with an application to the study of child survival in northeast Brazil. *Journal of the American Statistical Association* **92**: 426-435.

Schieke, T. H. and Jensen, T. K. (1997). A discrete survival model with random effects: an application to time to pregnancy. *Biometrics* **53**: 318-29.

Shah, A., Laird, N. and Schoenfeld, D. (1997). A random effects model for multiple characteristics with possibly missing data. *Journal of the American Statistical Association* **91**: 775-79.

Shun, Z. (1997). Another look at the salamander mating data: a modified laplace approximation approach,. *Journal of the American Statistical Association* **92**: 341-49.

Strand, S. (1997). Pupil progress during key stage 1: a value added analysis of school effects. *British Educational Research Journal* **23**: 471-88.

Thomas, S., sammons, P. and Mortimore, P. (1997). Differential Secondary School effectiveness: examining th esize, extent and consistency of school and departmental effects on GCSE outcomes for different groups of students over three years. *British Educational Research Journal* **23**: 451-469.

Thomas, S., sammons, P., Mortimore, P. and Smees, R. (1997). Differential secondary school effectiveness: comparing the performance of different pupil groups. *British Educational Research Journal* **23**: 451-70.

Thomas, S., Sammons, P., Mortimore, P. and Smees, R. (1997). Stability and consistency in secondary schools' effects on students GCSE outcomes over three years. *School effectiveness and school improvement.* **8**: 169-197.

Thum, Y. M. (1997). Hierarchical linear models for multivariate outcomes. *Journal of Educational and behavioural Statistics* **22**: 77-108.

Tymms, P., Merrell, C. and Henderson, B. (1997). The first year at school: a quantitative investigation of the attainment and progress of pupils. *Educational research and evaluation* **3**: 101-118.

Wright, D. B. (1997). Extra-binomial variation in multilevel logistic models with sparse structures. *British Journal of Mathematical and Statistical Psychology* **50**: 21-30.

Wulfsohn, M. S. and Tsiatis, A. A. (1997). A joint model for survival and longitudinal data measured with error. *Biometrics* **53**: 330-39.

Xue, X. and Brookmeyer, R. (1997). Regression analysis of discrete time survival data under heterogeneity. *Statistics in Medicine* **16**.

Yair, G. (1997). Teacher's polarisation in heterogeneous classrooms and the social distribution of achievement: and Israeli case study. *Teaching and teacher education* **13**: 279-293.

Zackin, R. and Wei, L. J. (1997). Analysis of repeated virological measurements based on cell dilution assays. *Statistics in Medicine* **16**: 571-82.

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

Workshops and Courses in Multilevel Modelling in 1997

Trondheim, Norway, 3-6 November 1997: This workshop was organised by the Department of Sociology and Political Science, Norwegian University of Science and Techonology (NTNU) in Trondheim as a part of the Department's Ph.D. program. The course was for beginners without extensive background in statistics. Attending the course and writing a paper applying multilevel analysis, give the participants 3 credits in the Ph.D program. Participants from other fields and/or from other parts of the country was invited to participate. The participants (12 in all) came from fields including Psychology and Veterinary sciences.

The workshop was administered by Kristen Ringdal from the Department of Sociology and Political Science and the main lecturers were Jon Rasbash and Min Yang from the Multilevel Models Project at the Institute of Education in London. The course was entirely based on the beta-version of *MLwiN*, the Windows 95/NT version of *MLn*. The program stood up well in the demonstrations and exercises over the four days. Its new graphical user interface makes it quite superior to the current (dos) version of *MLn*. *MLwiN* will obviously lower the costs of beginners starting to learn multilevel analysis. The experieneced used will also see the benefits, especially in the handling of non-linear, and multivariate models that both are more accessible through simplified procedures. Our impressions from the few days of intensive exercises are positive, although all functions of the final version were not implemented in early November

. (Professor *Kristen Ringdal*)

Postdam, Germany, 28-29 November 1997: This was a two-day workshop on Multilevel Analysis of Longitudinal Survey Data held at the Faculty of Economics, Business Management and Social Sciences, University of Potsdam (Germany). It was sponsored by the "Laengsschnitt-Werkstatt Berlin-Brandenburg" which is a recently established network of longitudinal researchers with institutional affiliations to universities and research centers in the Berlin-Potsdam area.

The workshop was on theory and methods of multilevel analysis and their applicability in longitudinal research. The list of speakers included social scientists and statisticians from Belgium, Germany, the Netherlands,the United Kingdom and Italy. All in all, 50 participants attended the workshop.

There were three invited presentations, each opening a major workshop section. The opening lecture on "*Multilevel analysis of longitudinal data: Models and issues*" was given by Joop Hox (Amsterdam / Utrecht) to introduce multilevel models for longitudinal data. A lecture on "*Crossed random effects in multilevel models*", with applications in social networks and spatial models was given by Tom A.B. Snijders (Groningen) to open the section on random cross-classified models. Min Yang (London) demonstrated the new *MLwin s*oftware with an example of longitudinal data analysis.

In addition the following papers were contributed:

- Regression Models for Clustered and Interdependent Observations by *Ulrich Poetter* and *Goetz Rohwer* (Bochum).
- Multilevel Models for Panel Data. Overview and Applications by *Uwe Blien* and *Katja Wolf* (Nuernberg).
- Random Effect Models for Event Data: a Study of University Dropout *by Massimo Montagni* and *Gori Enrico* (Florence).
- Longitudinal Meta-Analysis *by Cora Maas* and *Joop Hox* (Utrecht).
- Analysis of Product Usage Diary Data by *Helena Romaniuk, Philip Cooper* and *Chris Skinner* (Southampton).
- Application of Parametric and Nonparametric Hierarchical Modelling to Longitudinal Data by *A. Lopatatzidis, L. Moore, A.R. Cooper, T.J. Peters* (Bristol).
- Multilevel Analysis in Demographic Research: Some Critical Issues by *Giulia Rivellini* (Milano) and *Susanna Zaccarin* (Trieste).
- Analysing Unit-nonresponse in Panel Surveys with a Multilevel Cross Classified Model by *Jan Pickery, Geert Loosveldt* and *Ann Carton* (Leuven).
- Modelling Interaction Between Fertility and Working Careers in Italy: a Comparison Between Two Methodological Approaches by *Simona Drovandi* and *Carla Rampichini* (Firenze).
- Analyzing Change and Structural Effects by (Multivariate) Multilevel Analysis by *Uwe Engel* and *Manuela Poetschke* (Potsdam).

A selection of papers will appear in a book on Multilevel Analysis of Complex Survey Data (edited. by *Uwe Engel* and *Joop Hox*).

(Professor *Uwe Engel*)

<u>MLwiN/MLn Clinics in London 1998</u>

Tuesday February 10
Tuesday March 3
Tuesday April 7
Tuesday May 5
Tuesday June 2
at
Multilevel Models Project
11 Woburn Square, London WC1A 0SN
Contact Min Yang /Geoff Woodhouse for appointment

Tel: (0)171 612 6682 / 6657
Email: m.yang@ioe.ac.uk / teuegmw@ioe.ac.uk

Contribuors

We are grateful to the following who provided articles and news for this issue.

Uwe Engel, Dept. Social Structure Analysis, University of Potsdam (engel@rz.uni-potsdam.de)

Harvey Goldstein, Mathematical Sciences, Institute of Education (hgoldstn@ioe.ac.uk)

Enrico Gori, Dipartimento di Scienze Statistiche, University of Florence (gori@dss.uniud.it)

Andrew Jones, Department of Economics and Related Studies, University of York

Massimo Montagni, Dipartimento di Statistica, University of Florence (montagni@stat.ds.unifi.it)

Kristen Ringdal, Department of Sociology & Political Science, Norwegian University of Science & Technology (kristen.ringdal@sv.ntnu.no)

Nigel Rice, Centre for Health Economic, University of York (nr5@york.ac.uk)

Multilevel models where the random effects are correlated with the fixed predictors: a conditioned iterative generalised least squares estimator (CIGLS)

Nigel Rice[1], *Andrew Jones*[2] *and Harvey Goldstein*[3]

Introduction

This paper is a shortened version of a longer paper submitted for publication. It considers an extension to the iterative generalised least squares estimator (IGLS) to produce consistent estimates of fixed predictor parameters for multilevel models where the random effects are correlated with the fixed predictors and group sample sizes are small. The motivation for this work draws heavily on the econometrics literature on panel data estimators and the debate surrounding the use of fixed versus random effects.

There has been much debate in the literature on panel data estimators, between the two alternative specifications of fixed and random effects (see for example, Judge et al (1980), Hsiao (1986) and Baltagi (1995)). In its simplest form a variance components repeated measures or time series cross-sectional model can be specified in the following manner:

[1] Centre for Health Economics, University of York, York, Y01 5DD, England. E-mail: nr5@york.ac.uk
[2] Department of Economics and Related Studies, University of York, York, Y01 5DD, England.
[3] Institute of Education, University of London, London, WC1H 0AL, England.

Consider a simple 2-level variance components model. In the fixed effects model the level 2 group effects are treated as fixed but unknown to the observer. Such models allow the investigator to make inference conditional on the effects that are contained within the sample. In contrast, a random effects specification may be viewed as providing marginal or unconditional inference with respect to the population of all effects. It treats the effects as being random draws from an i.i.d distribution, typically Normal. The choice of specification may, in certain circumstances, be clear to the analyst and depend on the manner in which the data were sampled and the context of the investigation. For example, see discussions by Hausman (1978) and Goldstein (1995).

In this paper, we add to the debate on the relative merits of fixed and random effects by considering the general case of multilevel models. The procedure is general and may be extended more than two levels of the data hierarchy as well as random-coefficients and variable within group sample sizes.

Model specification

The 2-level variance components model may be written

$$y_{ij} = (\beta_X X)_{ij} + u_j + e_{ij}$$
$$i = 1,\ldots,N; \quad j = 1,\ldots,M \qquad (1)$$

Here X is a $K \times 1$ vector of exogenous variables, and β_X a $K \times 1$ vector of constants. We assume here that there are

M level 2 units or groups and N observations in total (and hence a total of N level 1 observations). Group sample sizes n_j are not required to be constant across the M groups. The components u_j and e_{ij} are residuals at level 2 and level 1 respectively, assumed to be i.i.d. with zero mean and constant variance:

$$\begin{aligned}
&\operatorname{cov}(u_j, u_{j'}) = \operatorname{cov}(e_{ij}, e_{i'j}) = 0,\\
&\operatorname{cov}(u_j, e_{ij}) = 0,\\
&E(u_j) = E(e_{ij}) = 0,\\
&\operatorname{var}(u_j) = \sigma_u^2,\\
&\operatorname{var}(e_{ij}) = \sigma_e^2.
\end{aligned}$$

The quantities of interest in (1) are the estimated parameters $\hat{\beta}_X$ (termed the fixed part parameters) and the estimated random components $\hat{\sigma}_u^2$ and $\hat{\sigma}_e^2$ (termed the random part parameters).

In the absence of correlation between the components of X and the level 2 random effects u_j, the iterative GLS estimator produces both efficient and consistent estimates of the fixed and random part parameters for fixed n_j (Goldstein, 1986). However where correlations between the level 2 random effects u_j and components of X exist, although IGLS estimation is efficient it is inconsistent as $M \to \infty$ when group sample sizes n_j are small (for example, see Blundell and Windmeijer (1997) for a full discussion relating to multilevel models).

Consistent estimation

Consistent estimation of β_X in (1) can be achieved by specifying the model as a fixed effects model and estimating by OLS. We write

$$y_{ij} = (\beta_X X)_{ij} + \sum_{i=1}^{N-1} \alpha_j d_{ij} + e_{ij}^* \quad (2)$$

If we pre-multiply (2) (whilst retaining the multilevel notation) by the idempotent matrix

$$Q = \{Q_j\}, \quad Q_j = I_{n_j} - J_{n_j} / n_j \quad (3)$$

where $\{\ \}$ denotes a matrix, and I_{n_j}, J_{n_j} are respectively the identity matrix and the square matrix of ones, of order n_j, we have (in matrix notation):

$$QY = QX\beta_X + QE^* \quad (4)$$

Applying OLS to (4) leads to consistent estimates of β_X. The estimator $(X^T QX)^{-1}(X^T QY)$ is known as the within groups or covariance estimator (CV) (for example, see Hsiao (1995)).

We now consider the following alternative conditioned iterative estimation procedure (CIGLS). IGLS estimation may be viewed as a two step procedure for each iteration. In the first step we re-express model (5) in matrix notation as:

$$Y = X\beta_X + S\beta_S + E$$
$$= Z\beta_Z + E \quad (5)$$

where $E = \{e_{ij}\}$ and

$$S = \{S_j\}, \quad S_j^T = (s_1, \ldots, s_M), \quad s_j = \overset{*}{s}_j \times \iota_{n_j},$$
$$\overset{*}{s}_j = \sum_{i=1}^{n_j} \hat{w}_{ij} / n_j,$$
$$\hat{W} = \{\hat{w}_{ij}\} = Y - X\hat{\beta}_X^* \quad (6)$$

where ι_{n_j} is a vector of ones of length n_j, $\hat{\beta}_X^*$ is the current estimate of β_X and S_j is obtained by stacking the vectors s_1 to s_M and is of length $\sum_{j=1}^{M} n_j = N$. In other words, the vector S consists of the group means of the estimated residuals from the previous iteration. Once S is constructed, updated estimates of β_X are then obtained through GLS estimation of (5).

In the second step, we condition on $\hat{\beta}_X$ and form the matrix $\overset{*}{Y} = \hat{W}\hat{W}'$. By stacking the columns of $\overset{*}{Y}$ these are regressed on the random parameter design matrix and GLS estimation produces the parameters of interest; $\hat{\sigma}_e^2$ and $\hat{\sigma}_u^2$ (for a variance components model, the random parameter design matrix is the block diagonal matrix leading to the covariance matrix V with elements $\sigma_e^2 I_{n_j} + \sigma_u^2 J_{n_j}$ for each group or block (for a full discussion, see Goldstein, 1995).

Suitable starting values for $\hat{\beta}_X^*$ may be obtained by OLS estimation of (2). Iteration of the two steps proceeds to convergence defined by a pre-assigned tolerence for $(\hat{\beta}_X - \hat{\beta}_X^*)$.

Convergence

We can re-express S in (5) as

$$S = (I - Q)(Y - X\hat{\beta}_X^*) \quad (7)$$

where Q is defined in (3) and I is the identity matrix. It then follows directly from (5) that

$$Y = X\beta_X + ((I - Q)(Y - X\hat{\beta}_X^*))\beta_S + E \quad (8)$$

If at convergence we have $\hat{\beta}_X = \hat{\beta}_X^*$, and $\hat{\beta}_S = 1.0$, (5) reduces to

$$QY = QX\beta_X + E \quad (9)$$

which is equivalent to the within groups specification (4) with $E = \overset{*}{E}Q.$.

It follows immediately therefore that the GLS estimator for the full set of fixed coefficients $\beta_Z = \begin{pmatrix} \beta_X \\ \beta_S \end{pmatrix}$ in (9), namely $(Z^T V_E^{-1} Z)^{-1}(Z^T V_E^{-1} Y)$, where V_E is the block diagonal covariance matrix $\left(= \begin{cases} \sigma_e^2 I_{n_j} + \sigma_u^2 J_{n_j}, & i = i', \\ 0, & i \neq i', \end{cases} \right)$, provides both the efficient and *consistent* (maximum likelihood under Normality) estimator of β_X.

Note, that in comparison to the GLS estimator, the OLS estimator ignores the lack of independence induced by premultiplying E in the equivalent multilevel fixed effects specification of (1) or (5) by Q. Also, if at convergence we obtain a value of $\hat{\beta}_S$ substantially different from 1.0 this may indicate misspecification in either the fixed or random parts of the model and could form a basis for diagnostic specification checks.

This procedure can be extended to the case where there are group level explanatory variables and where the LSDV is inapplicable. It also readily extends to the random coefficient case and to any number of levels. In the next section we look at some simulation results.

Simulations

We consider the simulation of a random coefficient at level 2, such that:

$$y_{ij} = 1 + 1x_{1ij} + 1.5x_{2ij} + v_j + \lambda_j x_{2ij} + e_{ij}$$

where

$$n_j = 5 \quad \forall j, \quad j = 1,\ldots,30$$
$$x_{1ij} \sim N(0,1), \quad x_{2ij} \sim N(0,1), \quad v_j \sim N(0,1),$$
$$\lambda_j \sim N(0,1.25), \quad e_{ij} \sim N(0,1.5)$$
$$\sigma_{x_{1ij}x_{2ij}} = 0, \quad \sigma_{x_{1ij},v_j} = 0.75, \quad \sigma_{x_{2ij},v_j} = 0,$$
$$\sigma_{e^*_{ij}v_j} = 0, \quad \sigma_{x_{1ij},\lambda_j} = 0.671, \quad \sigma_{\lambda_j,v_j} = 0.563.$$

The results of simulating the above model are presented in Table 1. We obtain improved estimates using CIGLS over IGLS, particularly for β_1 which corresponds to the LSDV estimate. The difference between the estimated constant and β_2 derived through LSDV and CIGLS is due to the different assumptions concerning the 'baseline group' adopted. For CIGLS the constant represents a weighted average over all level 2 units, as does the estimate for β_2. In contrast, LSDV estimates are made relative to a chosen 'baseline group' and as such can only be interpreted respective to a particular level 2 unit.

Conclusions

The iterative generalised least squares estimator conditioning on the mean level 2 effects (CIGLS) provides both efficient and consistent estimates of β_X when the random effects are correlated with one or more of the fixed predictors and group sample sizes, n_j, are small. Modifications to the standard IGLS estimation routine are trivial and computationally undemanding in the case where variance components models are considered. More elaborate estimation is required where a random coefficient is also correlated with a fixed predictor, but again this can be handled adequately using existing software (*MLn*, Rasbash, J. et al (1995)). In all cases, the procedure avoids the use of dummy variables (as in the standard LSDV estimator) and hence the associated loss in degrees of freedom, and the requirement to transform data to represent deviations from group means.

Table 1. Random coefficient model

	OLS LSDV			Multilevel IGLS			Multilevel CIGLS		
	Mean	SD	MSE	Mean	SD	MSE	Mean	SD	MSE
Random effects									
Level 2									
σ_v^2	--	--	--	0.966	0.373	0.140	1.029	0.389	0.152
σ_λ^2	--	--	--	1.243	0.448	0.266	1.253	0.449	0.262
$\sigma_{(\lambda,v)}$	--	--	--	0.537	0.304	0.380	0.564	0.312	0.415
Level 1									
σ_{e*}^2	2.251	0.342	0.180	1.502	0.227	0.299	1.494	0.225	0.307
Fixed predictors									
Constant	1.026	1.297	5.575	1.002	0.211	4.035	1.003	0.215	4.035
β_1	1.007	0.133	1.004	1.093	0.129	0.839	1.007	0.133	1.003
β_2	1.494	1.567	4.719	1.507	0.235	2.285	1.506	0.234	2.285

References

Baltagi, B.H. (1995). *Economic analysis of panel data*. New York, John Wiley & Sons.

Blundell, R. and Windmeijer, F. (1997). Correlated cluster effects and simultaneity in multilevel models. *Health Economics*; **1**: 6-13.

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

Goldstein, H. (1995). *Multilevel Statistical Models (second edition)*. London: Edward Arnold.

Hausman, J.A. (1978). Specification tests in econometrics, *Econometrica* **46**: 1251-71.

Hausman, J.A. Taylor, W.E. (1981). Panel data and unobservable individual effects. *Econometrica*, **49**: 1377-98.

Hsiao, C. (1995). *Analysis of Panel Data*. Econometric Society Monographs; Cambridge: Cambridge University Press.

Judge, G.G. Griffiths, W.E. Hill, R.C. Lee, T-C. (1980*). The theory and practice of econometrics*, New York, John Wiley & Sons.

Mundlak, Y. (1978). On the pooling of time series and cross section data, *Econometrica* **46**: 69-81.

Rasbash, J. and Woodhouse, G. (1995*): MLn: command reference guide.* London: Institute of Education.

Random effects models for event data ---- evaluating effectiveness of universities through the analysis of student careers

Enrico Gori1[1] - *Massimo Montagni*[2]

1. Introduction

The analysis of university student's careers by duration models has received increasing attention in recent years (Zwick et Braun 1988; Civian 1990). There are many reasons for applying this kind of analysis to the Italian University. First of all the average time to graduation (7-8 years) is much higher than that institutionally planned (4-5 years). Several students are still enrolled after 17 years. Finally a high percentage of freshmen (25-35%, depending on the faculty) dropout in the first two years, while others (20-30%) dropout after 3 or more years, having taken a small part of the planned exams. In this paper the main concern is with dropout problem. Although a competing risk model should be used, we confine ourselves to the analysis of dropout risk, by censoring all graduate students at the beginning of the interval in which this event occurred (Allison 1982).

2. Evaluating effectiveness in education

Following Willms (1992), in defining effectiveness it is worth to distinguish between:

- type A effects, which include the effects of institutional policy and practices (e.g. resources), characteristics (e.g. class sizes), composition (e.g. socioeconomic background of pupils), and the effects of social and economic factors (e.g. local employment rate);
- type B effects, which includes only institutional policy and practice, controlling for the other factors mentioned above.

In both cases we assume, obviously, to control for individual inputs. With reference to the basic variance components model:

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

where i indexes the "agent" (school, faculty), and j indexes the individuals, if $\mathbf{x}_{ij}$ includes only individual covariates, u_i measures type A effects, which are of main interest for the "clients" (students and families) in order to choose the "best" agent. If we include in $\mathbf{x}_{ij}$ also the other factors (characteristics, composition, etc.), out of the control of agents and principals, u_i can be interpreted as type B effects, and they are of main interest for "principals" (administrators) (*Goldstein* and *Thomas* 1996). But in order to get deeper insight into type B effects, it seems to be appropriate to distinguish between:

[1] Dipartimento di Scienze Statistiche, via Treppo 18, 33100 Udine. Email: gori@dss.uniud.it

[2] Dottorato in Statistica Applicata - Dipartimento di Statistica 'G.Parenti', via Morgagni 59, 50136 Firenze. Email:montagni@stat.ds.unifi.it

- B1 effects, which are related to observable characteristics of the "production function" of each agent, under its control (for example the amount of resources employed, the class size etc.) Note: *we think that these effects should include also what Willms defines as "characteristics (e.g. class sizes)" if they are under the control of the agent or principal*.

- B2 effects, which are related to unobservable management ability of the agent.

A possible way to identify these effects is to specify the overall type B effects as a function of observable characteristics of the production process at the "agent" level:

$$u_i = h(\mathbf{z}_i;\gamma) + v_i$$

where $\mathbf{z}_i$ are agent specific observable characteristics, $h(.)$ is a "technological" relation, γ are parameters to be estimated; $h(\mathbf{z}_i;\gamma)$ will measure type B1 effects, while v_i will measure type B2 effects. Both effects are of interest for the "principal", but B1 effects are more important to plan "production strategies", while B2 effects are more useful to implement "incentives policies" (*Laffont* and *Tirole* 1993) aimed to growth effectiveness and stimulate competition between agents. For reasons of equity, principal cannot compare performances (effectiveness) of agents which differ for the amount of resources. Especially if resources are allocated between agents by the principal, and the allocation is based on effectiveness measures: in this case B2 measure should be used.

So it can be quite interesting to investigate the relation between output measured at the student level and resources, in particular the number of teachers with respect to students. Although *Aitkin* and *Longford* (1986) have shown that type B effects explain some significant portion of the individual variance, other studies do not support an effect of resources on student output (*Hanushek*, 1986; *Pincus* and *Rolph*, 1986). It must be noticed, however, that these last studies are based on data aggregated at school level.

3. Data and variables

A sample of 2400 freshmen, enrolled at one of the 11 faculties of the University of Florence, between years 1975-1984, was drawn from administrative archives. The follow-up period for each freshman is 8 years long. The time unit for events is the academic year. Students migrating to other universities are considered as censored observations, under the hypothesis that migration and dropout are independent. The reconstruction of the data for the analysis, with the method of failure indicators, provided a total of about 10500 student year-observations.

For each student were also available information, $\mathbf{x}_{ij}$, such as sex, cohort, region of residence, regularity in the school career preceding enrollment, eventual delay in the enrollment, kind of high school, score achieved at high school final examination-average sore of peers. For what concerns faculty (agent) characteristics $\mathbf{z}_i$, the number of freshmen and teachers were considered: these are time varying explanatory variables. For every year, the number of freshmen approximates the number of competitors that each student must face in the access to resources, i.e. teachers.

4. Model

Given the kind of data a discrete time duration model was formulated. The hazard function, conditional on student characteristics, resources and faculty, is defined as the probability:

$$\lambda_{ijt} = \lambda(t; \mathbf{x}_{ij}, z_{it}, v_{it}) = P(T = t | T \geq t, \mathbf{x}_{ij}, z_{it}, v_{it})$$

If we choose for the hazard a logistic specification, a multilevel discrete time duration model can be specified as follows:

$$(y_{ijt} | \lambda_{ijt})_{i,j,t} \overset{independent}{\sim} B(1, \lambda_{ijt})$$
$$\text{logit}\lambda_{ijt} = \alpha_t + \mathbf{x}'_{ij}\beta + \mathbf{z}'_{it}\gamma + v_{it}$$

with $i = 1 \ldots N_2$, the number of faculties, $t = 1 \ldots k$, the number of intervals (8 years), $j = 1 \ldots n_{it}$, the number of students enrolled in faculty i which are at risk in the interval t. Baseline hazard parameters are assumed to be random at faculty level and across time intervals. The agent variables are introduced by a translog function. We chose to specify the model in a Bayesian framework, with non informative prior for the parameters, that is:

$$\{(\alpha_t^{(r)})_{t=1\cdots k}, \beta, \gamma\} \propto uniform,$$
$$(v_{it})_{i=1\cdots N_2, t=1\cdots k} \overset{i.i.d.}{\sim} N(0, \sigma_v^2),$$
$$\sigma_v^2 \sim IG(0.001, 0.001).$$

The reason for choosing a Bayesian approach was mainly motivated by the small number of second level units (11 faculties). We estimated the model with Monte Carlo Markov Chain techniques; we set up a Metropolis algorithm (Tierney 1994).

5. Empirical results

A summary of the results is presented in table 1. For what concerns type B1 effects, although most of the parameters involving freshmen have a posterior distribution that includes zero, the number of teachers is significant in explaining dropout hazard. Moreover, if we look at the graph of the dropout hazard in the first year (fig. 1) for the baseline student, this is an increasing function of the number of freshmen and a decreasing function of the number of teachers. This means that, although number of freshmen is not significant, the sign is that expected. It seems that a deeper analysis is needed to identify the "right" number of students in competition for access to resources. The posterior expectation of standard deviation of faculties effects is substantially larger than 0 when compared to the standard error, this means that, controlling for B1 effects, "ability of the faculties" (B2 effects), matter in explaining individual output. Note *it may be questionable to define as "more effective" a faculty with a lower dropout hazard. But, on one hand dropout rates of Italian universities are so high that it is* *very important to try to reduce it; on the other hand we can try to correct possible distortions of this indicator, using other indicators, such as the success of graduates in the labor market.* We derived also simultaneous interval for ranks of faculties, based on probability of dropout within the first three years. From figure 2 we can see that there are 2 groups of faculties for which 90th interval for ranks do not overlap.

6. Conclusions

In evaluating effectiveness of public services such as education *Willms* (1992) introduced the distinction between type *A* and type B effects. In this work we introduce a distinction of type B effects in B1 and B2 effects. The results of the analysis show that in this case study this classification it's quite important. The distinction between the two types of effects is relevant for the principal, because B1 effects are more important to plan "production strategies", while type B2 effects are more useful to implement "incentives policies". This aspect is particularly important in situations such as Italian University, where resources are allocated by principal (government) to agents (universities), on the basis of their effectiveness (relative performances). Not doing so it is equivalent to compare schools performances without adjusting for characteristics of their students. This is not only illogical, but, in the case at hand, also iniquitous given the very large differences in the amount of resources per student, allocated by government between universities and faculties in Italy.

References

Aitkin M. and Longford N. (1986). Statistical modeling issues in school effectiveness studies, *JRSS (A)*, **149**: 1-43.

Allison P.D. (1982). *Discrete-time methods for the analysis of event histories*, *Sociological Methodology*, Leindhart S. ed., S. Francisco, 61-98.

Civian J.T. (1990). *Using Proportional Hazard Models to Examine Time to Doctorate*, PhD Thesis, Haward University, Cambridge, Ma, USA.

Goldstein H. and Spiegelhalter D.J. (1996). League tables and their limitations: statistical issues in comparisons of institutional performance, *JRSS (A)*, **159** (3): 385-443.

Goldstein H. and Thomas S. (1996). Using examination results as indicators of school ad college performance, *JRSS (A)*, **159** (1): 149-163.

Hanushek E. A. (1979). Conceptual and empirical issues in the estimation of educational production function, *Journal of Human Resources*, **14** (3), 351-388.

Laffont J.J and Tirole J (1993) *A Theory of Incentives in Procurement and Regulation*, Cambridge, MIT Press.

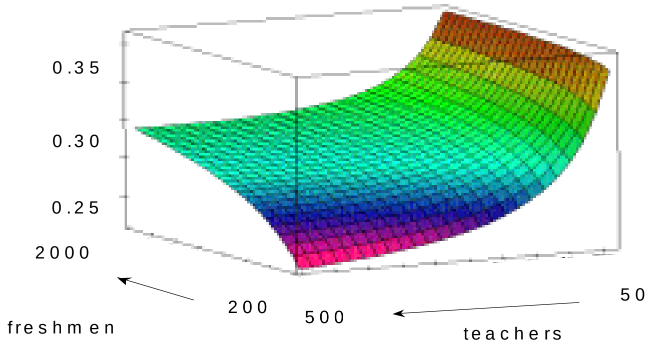
Pincus J and Rolph J. E. (1986) *How Much Is Enough? Applying Regression to a School Finance Case*, in Statistics and the Law, DeGroot M. H., Fienberg S. E., Kadane J. B., John Wiley, New York.

Tierney L. (1994) Markov Chains for exploring posterior distributions, *The Annals of Statistics*, No. (4), 1701-1762.

Willms J. D. (1992) *Monitoring School Performance: A Guide for Educators*, London, Falmer.

Zwick R. and Braun H. I. (1988) *Methods for analyzing the attainment of graduate school milestones: A case study*, GRE Board Professional Report No. 86-3P; ETS Research Report No. 88-30, Princeton, NJ: Educational Testing Service, USA.

Fig. 1 - Dropout hazard in the first year as function of number of teachers and freshmen



Tab. 1 - Posterior distributions from MCMC based on iterations 30,000-50,000 (samples of 4,000 values)

Parameter	Mean	S.D.	5th cen.	95th cen.
Baseline parameters				
α_1	-2.27	0.22	-2.63	-1.94
α_2	-3.08	0.22	-3.44	-2.72
α_3	-3.64	0.24	-4.06	-3.27
α_4	-3.82	0.23	-4.24	-3.43
α_5	-4.32	0.27	-4.77	-3.89
α_6	-3.98	0.26	-4.42	-3.56
α_7	-4.29	0.30	-4.78	-3.80
α_8	-3.98	0.30	-4.47	-3.50
School career preceding enrollment				
Regular	-----			
Not regular	0.28	0.11	0.09	0.44
Delay in the enrollment				
No	-----			
Yes	0.89	0.08	0.76	1.03
Region of residence				
Florence	-----			
Tuscany	-0.48	0.11	-0.66	-0.31
Rest of Italy	0.11	0.09	-0.04	0.25
Kind of high school				
Classical studies	-----			
Scientific studies	0.29	0.14	0.12	0.47
Techniques studies	1.47	0.11	1.31	1.66
Other	1.21	0.10	0.97	1.44

Parameter	Mean	S.D.	5th cen.	95th cen.
Sex				
Male	-----			
Female	0.03	0.09	-0.11	0.18
High school score - average score	-0.041	0.005	-0.05	-0.03
Cohort				
1975	-----			
1976	0.28	0.17	-0.002	0.57
1977	0.18	0.17	-0.09	0.46
1978	0.30	0.16	0.03	0.57
1979	0.50	0.06	0.23	0.77
1980	0.07	0.20	-0.27	0.41
1981	-0.03	0.21	-0.39	0.33
1982	0.11	0.17	-0.17	0.39
1983	0.09	0.16	-0.17	0.36
1984	-0.04	0.17	-0.31	0.24
Institutional factors				
Log *Teachers*	-0.52	0.11	-0.70	-0.34
(Log *Teachers*)^2	-0.04	0.26	-0.47	0.39
Log *freshmen*	-0.02	0.13	-0.22	0.20
(Log *Freshmen*)^2	-0.23	0.21	-0.59	0.11
(Log *Teachers*)*(Log *Freshmen*)	-0.17	0.25	-0.60	0.24
σ_v	0.28	0.06	0.18	0.38

Fig. 2 - 90 percent interval for simultaneous ranks based on dropout probability within first three years