DOCUMENT RESUME

ED 445 630 HE 033 303

AUTHOR	Patrick, William J.
TITLE	Estimating First-Year Student Attrition Rates: An Application of Multilevel Modeling Using Categorical Variables.
PUB DATE	2000-05-00
NOTE	25p.; Paper presented at the Annual Meeting of the Association for Institutional Research (40th, Cincinnati, OH, May 21-24, 2000).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Academic Persistence; College Admission; *College Freshmen; Dropout Rate; *Dropout Research; Educational Research; Foreign Countries; Higher Education; Research Methodology; School Holding Power; *Student Attrition; Withdrawal (Education)
IDENTIFIERS	*AIR Forum; Multilevel Analysis

ABSTRACT

This study examined first-year attrition at a large, urban university in the United Kingdom, demonstrating the application of multilevel modeling to the issue of student attrition. A sample of 2,679 full-time, first-year students studying the 20 most common subject areas was identified. Students were divided into four groups depending on their entry route: entering after the 6th year of secondary school after obtaining sufficient Scottish school-leaving qualifications to enter the university in the 5th year; entering with English school-leaving qualifications (which require 6 years of secondary school); entering with Scottish school-leaving qualifications obtained in the year of their leaving secondary school (5th or 6th year); and entering with non-school-leaving qualifications, mainly gained at college. Students' subsequent first-year dropout rates were, respectively, 5.6 percent, 7.6 percent, 11.6 percent, and 18.2 percent. A multilevel random coefficient model was fitted to the data. Data from logistic regression analysis and multilevel analysis indicated that students studying certain subjects had significantly different withdrawal rates. There was little evidence of an interaction effect between subjects and entry routes. There were some significant differences among the withdrawal rates of students in the four entry route groups. (Contains 30 references.) (SM)

Estimating First-Year Student Attrition Rates:
An Application of Multilevel Modeling Using Categorical Variables

William J. Patrick
Planning Officer
University of Glasgow
University Avenue
Glasgow G12 8QQ
Scotland
+44-141-330-4218
Fax: +44-141-330-4920
E-mail: W.Patrick@admin.gla.ac.uk

Paper presented at the 40th Annual Forum of
The Association for Institutional Research
Cincinnati, Ohio
May 21 – 24, 2000

Estimating First-Year Student Attrition Rates: An Application of Multilevel Modeling Using Categorical Variables

Abstract

There have recently been significant theoretical developments in multilevel statistical modeling, and improved software is readily available. This study demonstrates the application of multilevel modeling to one of the most common issues that confront institutional researchers: that of student attrition, where the response variable is typically binary rather than continuous. Comparisons are made with a traditional logistic regression approach. The data pertain to one large university. The techniques illustrated may be extended to the analysis of data sets encompassing many institutions, making meaningful inter-institutional comparisons of performance feasible even when there is hierarchical clustering present in the data.

Introduction

This study focuses on first-year retention in a large urban university. Results using multilevel statistical modeling are compared with those obtained using conventional logistic regression techniques. First-year student dropout rates form the binary dependent variable. There are only two explanatory variables in the example illustrated: individual students' entry qualifications, categorized into one of four entry qualification routes, and the subjects studied in first year. Each student's withdrawal can be expected to be influenced to a certain extent by her or his own entry route, but not by that of other students. On the other hand, it is conceivable that there is a clustering effect in the non-continuation rates of students studying the same subjects.

The best established and most tested theory of student attrition is Tinto's Student Integration Model (Tinto, 1975, 1987 and 1993). It is characterized by the view that student departure is a consequence of the interaction between the individual student and the college or university as an organization. It consequently offers the attractive possibility of effective intervention on the part of college and university authorities in order to ameliorate withdrawal rates. Empirical evidence indicates "partial support for the main elements in Tinto's model" (Braxton et al., 1997). The purpose of this study is neither to support nor to refute any aspects of the model, but to illustrate some of the apparent anomalies that can occur when conducting quantitative investigations of student attrition at different levels of aggregation, and to demonstrate techniques that can reveal hierarchical structures in the data.

The Higher Education Funding Council for England's comprehensive analysis (HEFCE, 1999a and 1999b) of 175 higher education institutions in the United Kingdom demonstrated a clear relationship between entry qualifications and non-completion, particularly in students aged under 21 on entry. The relationship has been examined in more detail in the United States. In his study of Bachelor's Degree attainment, Adelman (1999) emphasizes the significance of a broader concept of the "academic resources" that students bring to higher education. Within this composite measure curriculum contents are more important than both test scores and class rank/GPA.

Writing about the U.K. experience, Yorke (1999) has identified three primary causes of withdrawal among full-time students: a mismatch between students and their choice of field of study, financial difficulties, and poor quality of the student experience. This latter factor refers to the "quality of the teaching, the level of support given by staff and the organization of the program". Evidence of the importance of the first of these three factors is relatively strong (Thomas et al., 1996; Ozga and Sukhnandan, 1997). The concept of 'matching' between students and the subjects they study forms the focus of this study. Are there particular subjects or groups of subjects within a given institution that have higher attrition rates than others, even allowing for the fact that they attract students from different academic backgrounds in differing proportions? Are there some subject groups in which discrepancies among the dropout rates of students using different entry routes are less pronounced? Finally, do entry routes have a significant effect on dropout rates? These are the three issues to be explored.

Statistical Considerations

It is hypothesised that there are likely to be greater similarities in the performance of students studying any particular subject or group of subjects in a large institution than there are between students studying different subjects. This is because, regardless of their background attributes and abilities, students studying different subjects are exposed to different teaching methods and different levels of academic support. Different academic departments support different subjects. Throughout this paper first-year subjects are treated as being synonymous with the departments in which they are taught. In certain areas, teaching may serve simply to provide a general overview of and introduction to a subject, while in others the intention may be to explain detailed concepts and perhaps to focus on particular aspects with which students can be expected to experience difficulty. Laboratory sessions may be considered necessary for most science-based subjects; on the arts side, laboratories are much less common. The frequency and size of tutorial meetings can also vary among subjects. Remedial tutoring, personal help and counseling may be more readily accessible in some areas than in others. In short, students are contextualized by the departments teaching the subjects which they study.

The clustering of students into contextual hierarchies causes particular statistical difficulties. One of the assumptions that underlies ordinary least squares (OLS) regression analysis is that the stochastic error terms are independent of one another. Clustering imposes a correlation structure on the data, and the use of OLS techniques will cause the standard errors to be underestimated. This in turn will cause confidence intervals to be too narrow and significance tests to reject the null hypotheses when, in fact, they should not.

A common concern in dealing with hierarchical data is the danger of incorrectly inferring to individuals the associations that have been observed at the aggregate level. Robinson's study (1950) of the relationship between literacy and ethnic background provides an early example of what has become known as the ecological or aggregation fallacy. He found a correlation of 0.95 at the level of census divisions between mean literacy rates and mean proportions of Black Americans. This was in marked contrast to the correlation of 0.20 at the individual level. The more general issue is to determine the most appropriate level at which to conduct one's analysis. The "unit of analysis" problem is well recognized as a potential source of divergent research findings when considering the effects of college characteristics on student development (Pascarella and Terenzini, 1991, pp. 682-7; Ethington, 1997). Tinto (1993, p. 33) has alluded to the "complex interplay of individual and institutional forces which shape the extent and patterning of student departures from higher education". Multilevel modeling is ideally suited to examining the association between individuals' characteristics and the contextual settings to which they belong, because it is possible to partition variation in the response variable into that attributable to factors associated with different levels in a contextual hierarchy, thereby avoiding the potential pitfalls of the ecological fallacy. Complex models can be created in which there are several variables at each level of in the hierarchy.

Creating a separate model for students within each departmental group precludes the possibility of making any inferences about the effects of such groups generally. By treating departmental groups as having random effects, multilevel modeling makes it possible to estimate each group's effect not only on the basis of data from that particular group (which may be sparse, and therefore subject to considerable sampling variation) but additionally by taking account of the effects of other groups and the characteristics which they have in common. When departmental groups are treated as observations drawn from a larger population of such groups, they are said to "borrow strength" from one another (Braun, 1989). A further advantage when the number of observations in each group is small is that the estimation methods used in multilevel modeling have the effect of pulling estimated group effects towards the overall mean. This general phenomenon is known as "shrinkage" (O'Hagan, 1994).

Conventionally, group effects are allowed for by using a series of dichotomous dummy variables, with the effect of individual departmental groups being estimated relative to a chosen baseline group. This involves estimating possibly many more coefficients than the alternative multilevel procedure. It can be difficult to assess interactions between the effects of the dummy variables and those of the other explanatory variables in the model. Traditional methods can produce misleading results when the number of observations in each group is small. Individual observations can exert undue influence on the estimated coefficients, leading to unsound statistical inferences.

Multilevel Estimation Procedures

In recent years there have been significant theoretical developments in multilevel statistical modeling and improved software has been released. Kreft and De Leeuw (1998) and Snijders and Bosker (1999) have provided useful introductory texts. Ethington (1997) has demonstrated the application of multilevel modeling to institutional research in higher education. The basic linear multilevel model has been extended to cover multivariate analysis, nonlinear data, repeated measures data, discrete response data, cross classification data and event history models (Goldstein, 1995). MLwiN software (Goldstein, Rasbash, Plewis, et al., 1998) has been developed from earlier MLn software to allow these types of model to be constructed using a graphical user interface, thereby making multilevel techniques more readily accessible to researchers than before.

Multilevel binary response models are a special case of hierarchical generalized linear models, for which parameter estimation is more difficult than in hierarchical linear models. Generally, a first- or second-order Taylor expansion of the link function is required. Marginal quasi-likelihood (MQL) methods involve approximations around the estimated fixed part of the model, whereas penalized or predictive quasi-likelihood (PQL) methods produce approximations around the fixed plus the random part of the model (Goldstein, 1991; Breslow and Clayton, 1993; Goldstein and Rasbash, 1996). The first-order MQL and PQL estimates of the variance parameters tend to have a downward bias (Rodríguez and Goldman, 1995). On the other hand, second-order MQL and PQL methods produce parameter estimates with less bias but a higher mean square error. The algorithms may not always converge to a solution. Small group sizes and zero cells can cause instability. First-order MQL is the most robust of the first- and second-order MQL and PQL methods.

For Bayesian modeling MLwiN can be used to implement Markov Chain Monte Carlo (MCMC) methods using either Gibbs sampling or Metropolis-Hastings procedures. The latter is appropriate for binary response variables. Generalized linear models with binomial error terms are supported. In Bayesian formulations prior information about the fixed or random parameters are combined with the data. These parameters are treated as random variables described by probability distributions, known as prior distributions. The distributions which result from the estimation procedure, which combines the prior distributions with the data, are know as posterior distributions (posteriors). MCMC methods approximate these distributions by making a large number of random simulated draws from the posteriors and summarizing them in terms of summary statistics such as means and standard deviations. In small samples MCMC takes account of the uncertainty associated with the estimates of the random parameters and can provide exact measures of the uncertainty. The quasi-likelihood methods tend to overestimate precision because they ignore this uncertainty. This is particularly important when obtaining posterior estimates for the residuals. More information on MCMC methods is given by Gilks et al. (1996).

An Application

The university studied is a large urban university in the U.K. that offers a wide range of academic subjects. In some areas the choice and content of courses is influenced by the requirements of a particular professional body, such as in medicine, dentistry, veterinary science, engineering, law and accountancy. In other areas, such as science, the arts and social sciences, students are able to choose subjects from a wide range. Information concerning the reasons for students' withdrawal is sparse. Students who were present for at least part of the 1998-99 academic year but who did not enrol again for the subsequent session are counted as dropouts in this investigation.

A sample of 2,679 full-time, first-year students studying the 20 most common subject combinations, predetermined or selected, were identified for analysis. The numbers of students in the corresponding 20 departmental groups ranged from 55 to 427, with dropout rates varying from 0% to 19.6% and with an overall average of 10.4%.

Although it is common in the U.K. to represent prior academic achievements in the form of a point score, there are various methodological disadvantages with this approach. For example, point scores obtained in closely allied subjects should arguably be given less weight than in cases where qualifications have been obtained in quite different subjects, although this is not common practice. For the purpose of the current study, students were instead split into four groups, according to their entry route: (1) 855 students in the sample came to the University from the sixth year of secondary school, having obtained sufficient Scottish school-leaving qualifications to enter the University in the fifth year ("Yr6_q5th"); (2) 439 students came to the University with English school-leaving qualifications, which require spending six years at secondary school (General Certificate of Education or "GCE"); (3) 921 students came to the University with Scottish school-leaving qualifications obtained only in the year of their leaving secondary school, either fifth year or sixth year ("Yr5Yr6q6"); and (4) 464 gained entry with non-school-leaving qualifications, mainly gained at college ("Other"). Their subsequent first-year dropout rates were 5.6%, 7.6%, 11.6% and 18.2%, respectively. Superficially, it appears that these are significantly different withdrawal rates, at least at the aggregate level. (Chi-squared = 51.1 with 3 degrees of freedom, which is highly significant.) Further investigation is required to establish whether similar differences exist also in different subject groupings within the institution.

Variance Between and Within Groups

It is important first to determine whether the data exhibit clustering effects. The existence of a hierarchical structure points to the appropriateness of multilevel modeling. The data were examined using the approach described by Snijders and Bosker (1999). Only the group-level covariate is taken into account at this stage.

In the case of a binary response variable, a slight modification of the chi-squared test can be used to give a broad indication of the existence of significant differences between groups. The conventional test statistic is in the form

$$\sum (O-E)^2 / E$$

where O and E represent the observed and expected counts respectively in a contingency table. With a binary response variable the test statistic is

$$X^2 = \sum_{j=1}^{N} n_j \frac{(p_j - P)^2}{P(1-P)}$$

where

N = number of (level-2) groups

n_j = number of observations in the jth group

p_j = probability of dropout in the jth group

P = overall probability of dropout

In this example $X^2 = 77.2$ with 19 degrees of freedom. This apparently highly significant result needs to be treated with caution because of the paucity of data in certain cells (Agresti, 1996).

An estimate of the population between-group variance, $\hat{\tau}^2$, may be calculated as the observed between-group variance less the contribution that within-group variance makes, on average, to the observed between-group variance. Algebraically

$$\hat{\tau}^2 = s^2_{between} - \frac{s^2_{within}}{\tilde{n}}$$

where

$$\tilde{n} = \frac{1}{N-1}\left\{M - \frac{\sum_j n_j^2}{M}\right\} \quad \text{and} \quad M = \sum_j n_j$$

In effect, $\tilde{n}$ is an adjustment to allow for the fact that not all groups contain the same number of observations. Under these circumstances the population between-group variance cannot be estimated solely from the observed between-group variance: the latter also includes some observed within-group variance (Searle, Casella and McCulloch, 1992).

We have

$$s^2_{between} = \left(\sum n_j (p_j - p)^2\right) \Bigg/ \left(M - \frac{\sum n_j^2}{M}\right)$$

and

$$s^2_{within} = \frac{1}{(M-N)} \sum_{j=1}^{N} n_j p_j (1 - p_j)$$

In this example, P = 0.104, $\hat{\tau}^2 = 0.0022$ and $\hat{\tau} = 0.0469$. The standard deviation is quite large, relative to the overall probability of dropping out, and this is suggestive of the existence of a hierarchical structure in the data. The subsequent multilevel analysis confirms this.

These are so-called analysis of variance (ANOVA) estimators in an "empty" two-level model in which the group-dependent probabilities, p_j, are specified without taking any further explanatory variables into account.

Logistic Regression

Standard logistic regression was first applied. The probability of the ith student dropping out may be expressed as

$$\mathrm{P}(\mathrm{dropout}_i) = \frac{e^{\beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i}}}{1 + e^{\beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i}}} = \pi_i$$

where

β_1 is the constant term

β_2 is the coefficient associated with the departmental groups

x_{2i} is the value of an indicator variable denoting the ith student's departmental group

β_3 is the coefficient associated with the entry routes

x_{3i} is the value of an indicator variable denoting the ith student's entry route

The conventional logit link function may be written as

$$\mathrm{logit}(\pi_i) = \ln\{\pi_i / (1 - \pi_i)\} = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i}$$

The binary (0, 1) response variable is then

$$y_i = \pi_i + \varepsilon_i$$

where

$y_i \sim \mathrm{Binomial}(1, \pi_i)$ and

ε_i has mean zero and variance $\pi_i(1 - \pi_i)$.

The expression $\pi_i/(1-\pi_i)$ represents the odds of a student's withdrawal, and is the probability of its occurrence divided by the probability of its non-occurrence. In this case the odds ratio is the ratio of the odds when an indicator variable, x, is 1 (signifying group membership) to the odds when $x = 0$ (signifying non-group membership), and is equal to $exp(\beta)$. If the 95% confidence interval for the odds ratio includes the value one, this provides strong evidence that the null hypothesis that the probability of occurrence is the same as the cohort average cannot be rejected.

A deviation coding system was used to represent each of the 20 departmental groups, as well as each of the four entry routes. The results are shown in Table 1. A deviation coding scheme allows the effect of any particular departmental group to be compared to the simple average of all groups combined (Menard, 1995). Two runs of SPSS software using different reference groups allowed test statistics for all departmental groups and all entry routes readily to be obtained.

It will be seen that only two entry routes have estimated dropout rates significantly different from the average. One is higher (*↑) and the other is lower (*↓). By using an indicator coding scheme, instead, and choosing the entry route with the highest observed dropout rate ("Other"), it is also possible to demonstrate that each of the other three entry routes has a significantly lower dropout rate. Conversely, it is possible to demonstrate that three entry routes have a significantly higher dropout rate than the lowest ("Yr6_q5th"). The perceived differences in the dropout rates for the four entry routes are dependent on the form of the model fitted.

It will also be observed from Table 1 that three of the 20 groups have a probability of dropout which is significantly higher than the average (*↑), and one which is below (*↓).

Figure 1 illustrates simultaneous confidence limits for each of the departmental groups. Each of the error bars on this "centipede" chart is plotted as $exp(\hat{\beta}_2 \pm 1.4 * \text{s.e.})$[1]. This gives a pragmatic, pairwise comparison of the dropout rates of any two departmental groups. When two error bars overlap we cannot reject the hypothesis that the two groups concerned have the same dropout rate. It will be seen that most of the groups analyzed do not have significantly different dropout rates. There are, however, a few groups on the left of the diagram which have significantly lower dropout rates than those on the far right, and vice versa. Group 1 on the left-hand side is exceptional in having no observed withdrawals. It is not possible to calculate test statistics for this group using conventional methods.

Table 1. Logistic Regression – Parameter Estimates and Confidence Intervals

Variable	β	s.e.	df	Sig	$exp(\beta)$	95% CI for $exp(\beta)$ Lower	Upper	Sig
Entry Route			3	0.00				
Yr6_q5th (1)	-0.75	0.13	1	0.00	0.47	0.36	0.61	*↓
GCE (2)	-0.18	0.15	1	0.24	0.84	0.63	1.12	
Yr5Yr6q6 (3)	0.14	0.11	1	0.17	1.16	0.94	1.42	
Other (4)	0.79	0.11	1	0.00	2.19	1.76	2.74	*↑
Group			19	0.00				
Group(1)	-0.16	0.49	1	0.74	0.85	0.32	2.23	
Group(2)	0.26	0.42	1	0.54	1.29	0.56	2.96	
Group(3)	1.01	0.45	1	0.02	2.74	1.14	6.56	*↑
Group(4)	0.45	0.50	1	0.37	1.56	0.59	4.15	
Group(5)	-0.17	0.54	1	0.76	0.84	0.29	2.45	
Group(6)	0.42	0.44	1	0.34	1.52	0.65	3.55	
Group(7)	0.68	0.49	1	0.16	1.98	0.76	5.12	
Group(8)	1.03	0.36	1	0.00	2.81	1.39	5.68	*↑
Group(9)	0.58	0.40	1	0.14	1.79	0.82	3.91	
Group(10)	0.04	0.44	1	0.93	1.04	0.44	2.44	
Group(11)	-0.49	0.47	1	0.30	0.62	0.25	1.54	
Group(12)	-0.94	0.45	1	0.04	0.39	0.16	0.95	*↓
Group(13)	0.73	0.47	1	0.12	2.09	0.83	5.24	
Group(14)	0.37	0.35	1	0.29	1.45	0.73	2.87	
Group(15)	0.72	0.42	1	0.09	2.06	0.90	4.69	
Group(16)	0.87	0.40	1	0.03	2.38	1.09	5.19	*↑
Group(17)	0.05	0.40	1	0.91	1.05	0.48	2.31	
Group(18)	0.69	0.40	1	0.08	2.00	0.91	4.37	
Group(19)	-6.06	5.83	1	0.30	0.00	0.00	212.91	
Group(20)	-0.07	0.55	1	0.90	0.93	0.32	2.74	
Constant	-2.47	0.32	1	0.00				

Figure 1 illustrates simultaneous confidence limits for each of the departmental groups. Each of the error bars on this "centipede" chart is plotted as $exp(\hat{\beta}_2 \pm 1.4 * \text{s.e.})$[1]. This gives a pragmatic, pairwise comparison of the dropout rates of any two departmental groups. When two error bars overlap we cannot reject the hypothesis that the two groups concerned have the same dropout rate. It will be seen that most of the groups analyzed do not have significantly different dropout rates. There are, however, a few groups on the left of the diagram which have significantly lower dropout rates than those on the far right, and vice versa. Group 1 on the left-hand side is exceptional in having no observed withdrawals. It is not possible to calculate test statistics for this group using conventional methods.

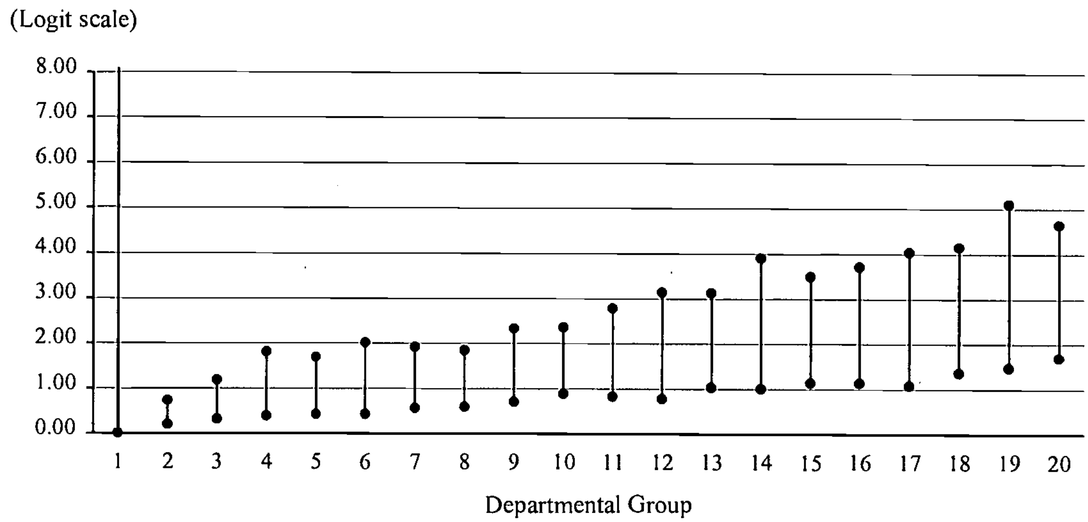

Figure 1. Logistic Regression – Simultaneous Confidence Intervals

Multilevel Models

A multilevel random coefficient model was then fitted to the data. The logit link function, which expresses the dependent variable as a linear function of the independent variables, was defined as

$$\text{logit}(\pi_{ij}) = \beta_{1j} + \beta_{2j}x_{2ij}$$

where

π_{ij} is the probability that the ith student in the jth departmental group drops out

β_{1j} is the intercept term in the model for the jth departmental group

β_{2j} is the slope term in the model for the jth departmental group

x_{2ij} is the average logit of the probability of dropout for the ijth student's entry route (= entrywgt$_{ij}$)

Defining x_{2ij} as the average logit of the probability of dropout for the ijth student's entry route allows the corresponding slope term, b_{2j}, to be fitted for each of the 20 departmental groups. The alternative would have been to have used four dichotomous dummy variables, as in the logistic regression model. It is advantageous to be able to represent each of the slope terms graphically although, in this case, the conclusions are essentially the same.

Both β_{1j} and β_{2j} are the sum of a fixed part and a random part of the model, the latter being assumed to vary among departmental groups with mean zero and constant variance. Formally:

$$\beta_{1j} = \beta_1 + u_{1j}$$
$$\beta_{2j} = \beta_2 + u_{2j}$$

where

$$\begin{bmatrix} u_{1j} \\ u_{2j} \end{bmatrix} \sim \mathrm{N}(0, \Omega_u) \quad : \quad \Omega_u = \begin{bmatrix} \sigma^2_{u1} & \\ \sigma_{u21} & \sigma^2_{u2} \end{bmatrix}$$

The assumed distribution of the response variable is the same as in the logistic regression model, with only the subscripts differing in order to denote the two-level hierarchical structure of the data:

$$y_{ij} = \pi_{ij} + \varepsilon_{0ij}$$

where

$y_{ij} \sim \mathrm{Binomial}(1, \pi_{ij})$ and

ε_{0ij} has mean zero and variance $\pi_{ij}(1 - \pi_{ij})$.

The fixed parts of the model are β_1 and β_2, while the random part is made up of three terms: u_{1j}, u_{2j} and e_{0ij}. Whereas β_1 and β_2 are both level-2 (group) variables, the random part of the model consists of two level-2 terms (u_{1j} and u_{2j}) and one level-1 (individual) term (e_{0ij}). The splitting of the random part of the model into different expressions in this way is one of the features which distinguishes multilevel modeling from ordinary least squares (OLS) estimation, where there is typically only one error term.

A multilevel variance components model was also fitted to the data. The logit link function was changed to

$$\mathrm{logit}(\pi_{ij}) = \beta_{1j} + \beta_2 x_{2ij}$$

where

π_{ij} is the probability that the ith student in the jth departmental group drops out

β_{1j} is the intercept term in the model for the jth departmental group

β_2 is the slope term in the model

x_{2ij} is the average logit of the probability of dropout for the ijth student's entry route (= $\mathrm{entrywgt}_{ij}$)

It will be observed that β_2 is no longer suffixed with a 'j'. This implies that β_2 is a parameter which is assumed to be the same for all departmental groups. Only β_{1j} is the sum of both a fixed and a random term, the latter again being assumed to vary among departmental groups with a normal distribution having mean zero and constant variance. This may be expressed as

$$\beta_{1j} = \beta_1 + u_{1j} \ : \ \left[u_{1j} \right] \sim \mathrm{N}(0, \sigma^2_{u1})$$

The formulation of the response variable is the same as in the random coefficient model.

MLwiN software was used to fit these models, using three different estimation techniques. Two similar approximate methods of inference were used: first-order marginal quasi-likelihood (MQL1) and second-order penalized or predictive quasi-likelihood (PQL2). Thirdly, Markov Chain Monte Carlo (MCMC) Bayesian estimation was used. The results are summarized in Tables 2 and 3.

Table 2. Random Coefficient Model Parameter Estimates and Standard Errors

Parameter	Estimate (s.e.) MQL1		PQL2		MCMC	
Fixed:						
β_1 (intercept)	-2.199	(0.129)	-2.338	(0.153)	-2.378	(0.205)
β_2 (entrywgt)	1.240	(0.225)	1.323	(0.246)	1.329	(0.332)
Random Level 2:						
σ^2_{u1} (intercept)	0.215	(0.103)	0.329	(0.146)	0.646	(0.409)
σ_{u21} (covariance)	-0.050	(0.129)	-0.098	(0.170)	-0.214	(0.458)
σ^2_{u2} (entrywgt)	0.332	(0.297)	0.432	(0.359)	1.223	(1.049)

Table 3. Variance Components Model Parameter Estimates and Standard Errors

Parameter	Estimate (s.e.) MQL1		PQL2		MCMC	
Fixed:						
β_1 (intercept)	-2.220	(0.130)	-2.336	(0.152)	-2.360	(0.175)
β_2 (entrywgt)	1.333	(0.173)	1.368	(0.180)	1.381	(0.181)
Random Level 2:						
σ_{u1} (intercept)	0.223	(0.101)	0.330	(0.138)	0.499	(0.271)

The downward bias of the MQL1 estimates of the variance parameters in both Table 2 and Table 3 is clear. The higher mean squares of the PQL2 parameter estimates are also evident. The greater standard errors of the fixed parameters and the considerably higher standard errors of the variance parameters obtained by MCMC are particularly striking.

In the random coefficient model all three estimation techniques produce a high estimated standard error of σ_{u2}^2, relative to the parameter estimate, suggesting that this parameter is not significantly different from zero, so that the simpler variance components model is more appropriate.

This conclusion is reinforced by inspection of Figure 2, which represents the random coefficient model, and which was produced using MLwiN software. It shows the estimated intercept (or constant) and slope terms for each of the 20 departmental groups. Simultaneous confidence intervals for each pair of estimates are also shown. As in the logistic regression model, the error bars represent the relevant point estimates plus or minus 1.4 times their standard errors. The vertical axes represent logit scales, centered around the average estimate in each case. The error bars for each of the departmental groups are shown in rank order of each of the two sets of estimates separately, so the ordering of the departmental groups may consequently be different in the two charts illustrated. Although not shown, the chart of the estimated intercept terms and simultaneous confidence intervals for each of the 20 departmental groups using the variance components model is very similar to that shown for the random coefficient model.

As in the logistic regression model, it will be observed that there are a few pairs of departmental groups that have significantly different overall dropout rates, as demonstrated by non-overlapping error bars for some of the 20 intercept terms. However, none of the slopes are significantly different from any of the others, which is consistent with the high mean square variance of σ_{u2}^2 above.

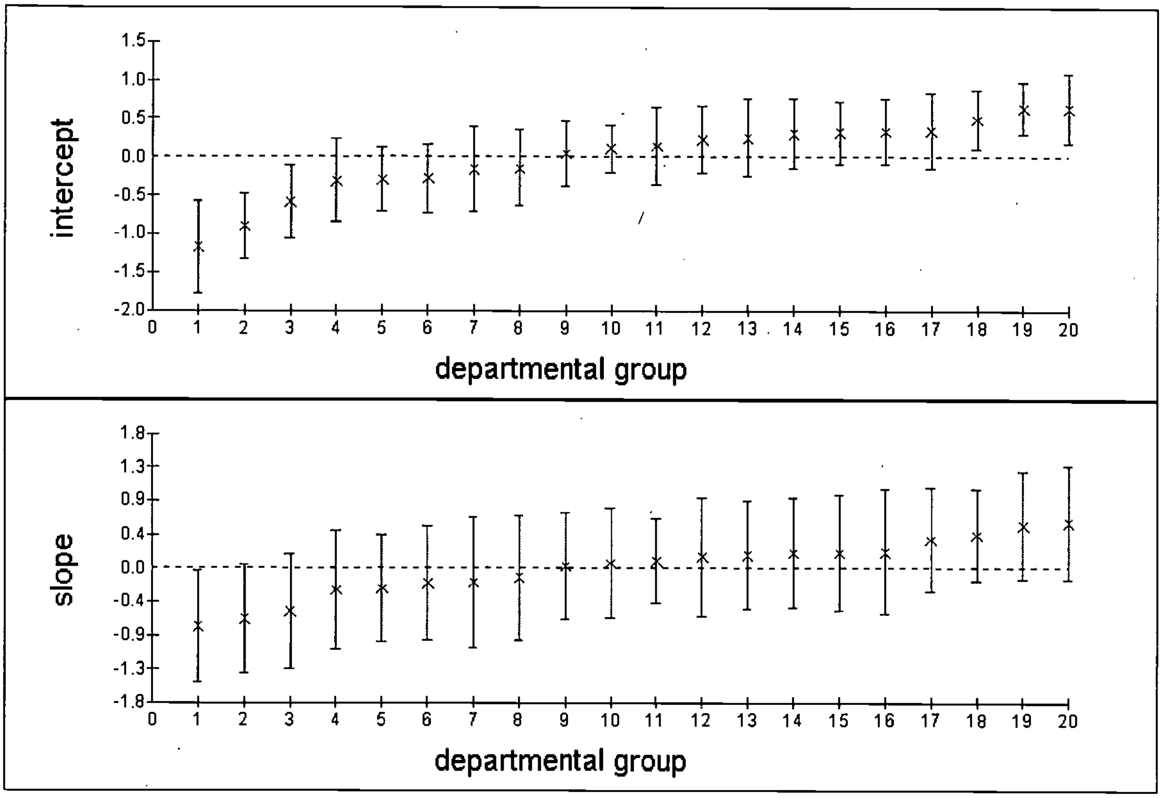

Figure 2. Random Coefficient Model – Simultaneous Confidence Intervals – Estimated Intercepts and Slopes ± 1.4 * s.e. (centered logit scales)

The fitted regression lines obtained using MQL2 estimation for the variance components model and random coefficient model the are shown in Figures 3 and 4 respectively. In each case the predicted value of the logit is expressed as a linear function of the four entry routes, each represented by its average logit of the probability of withdrawal. The 20 variance components lines are parallel, whereas the equivalent random coefficient model lines are not. This illustrates the crucial difference between a variance components model and a random coefficient model.

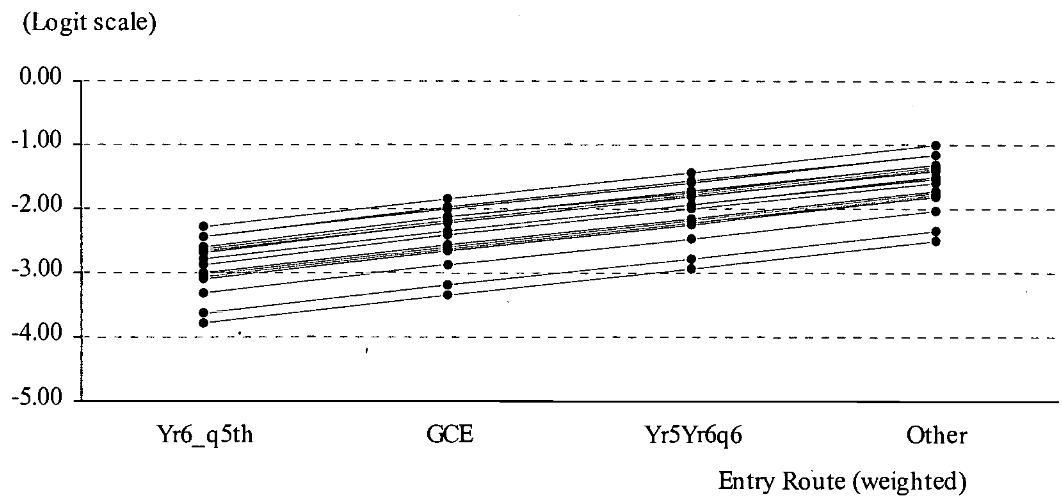

Figure 3. Prediction using a Variance Components Model

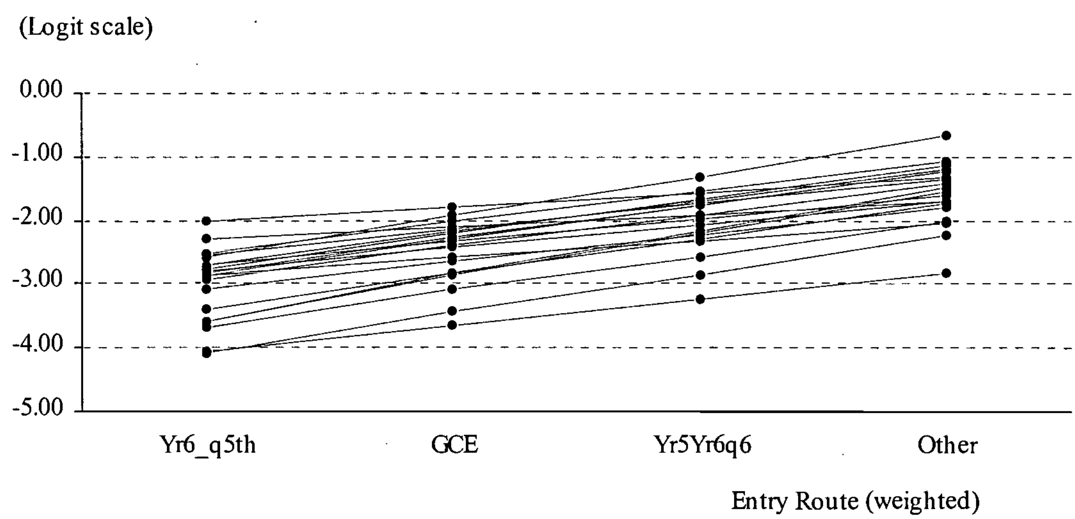

Figure 4. Prediction using a Random Coefficient Model

Conclusions

This example contains only two explanatory variables. Nevertheless, the study has practical value for the institution concerned, because certain subject areas have been identified in which attrition rates differ from the norm, even allowing for differing mixes in students' academic backgrounds. The results act as a valuable performance indicator, highlighting areas for further attention where positive lessons may be learned for dissemination across the institution. Care is needed in interpreting the results. There is likely to be confounding between student characteristics, the subjects they choose to study, and the characteristics of the particular academic environments in which they find themselves. This study is not sufficient on its own to establish the direction of causation: the initial or overriding reasons for differences in rates of attrition cannot be ascribed with any certainty to any one of these factors.

Of the three questions posed at the outset, one may be answered in the affirmative, one in the negative, while the answer to the third depends on the level of aggregation used.

First, there is evidence both from the logistic regression analysis and the multilevel analysis that students studying certain subjects do have significantly different withdrawal rates than most. This could be because the 'match' between students and the subjects they study is easier to achieve in certain areas than others, in accordance with Yorke's findings (1999) referred to above. An alternative explanation is that some subject groups are simply easier than others, perhaps because of lower academic standards, or perhaps because of a more supportive teaching and learning environment, which might in turn be a consequence of better departmental resources. It is also conceivable that there is a structuring within the student population that is not being captured in this analysis. For example, students studying vocationally oriented subjects may be more strongly motivated to persist than others. The likely existence of additional explanatory variables suggests that further investigation is merited.

Secondly, it will be seen that although the regression lines in Figure 4 for the random coefficient model are not parallel, there is not a great deal of variation in the point estimates of the slope coefficients shown in Figure 2. Such variation is not statistically significant, in any case. There is consequently little evidence of an interaction effect between subjects and entry routes. This suggests that there are no departmental groups that are more 'egalitarian' than others in their treatment of students having different academic backgrounds. The existence of slopes significantly different from others would perhaps have identified areas where remedial tutoring is either particularly effective or relatively weak.

Thirdly, it is possible to conclude from the logistic regression model that there were some significant differences among the withdrawal rates of the four entry qualification routes for the sample as a whole. However, the multilevel approach suggests that no significant differences can be detected when the data are analyzed at the level of individual departmental groups. This suggests that withdrawal rates would be little affected by changing the proportions of students admitted through different entry routes. It leaves open the possibility that the use of entrance criteria other than entry routes might prove more effective in reducing dropout rates. In particular, there may be lessons to be learned from the recruitment techniques employed by those departments having relatively low dropout rates and from the general characteristics and motivations of the students whom they recruit.

Further Implications and Limitations

The study serves as a warning to those who would seek to formulate league tables and make other comparisons based simply on headline, institution-level performance indicators. On closer investigation apparent variation in such figures may prove to be illusory and misleading (Goldstein and Spiegelhalter, 1996). Focusing on particular aspects of the organization will be more appropriate, at least from the perspective of an intending student, and is consistent with Tinto's emphasis on "Classrooms as Communities" (1997).

The example given relates only to one cohort of students. Clearly it would be desirable to amass more evidence by examining the performance of other cohorts. This could be achieved by adding a further layer to the multilevel hierarchy. The data relate only to one institution. It would be possible to further extend this type of analysis to make comparisons of different institutions' performance, using multilevel modeling to make proper allowance for the differing environments from which they attract students and variation in the characteristics and performance of the institutions as well as their students. Used with care, multilevel modeling is an important new tool in the statistical armory of the institutional researcher.

Note

1. A comparison between any pair of departmental groups may be made by constructing simultaneous confidence intervals of such a width that if they do not overlap then the null hypotheses that their two means are the same (H_0: $\mu_1 = \mu_2$) is rejected in favor of an alternative hypothesis, H_A: $\mu_2 > \mu_1$. If μ_1 and μ_2 are the means of two normal populations having the same variance with equivalent sample means of $\hat{\mu}_1$ and $\hat{\mu}_2$ and a pooled standard error of s_p, then H_0 is rejected when $\hat{\mu}_1$ and $\hat{\mu}_2$ are separated by more than twice a certain multiple (k) of the pooled standard error: $\hat{\mu}_2 - ks_p > \hat{\mu}_1 + ks_p$, or $\hat{\mu}_2 - \hat{\mu}_1 > 2ks_p$. Because the two populations are assumed to be normally distributed, the 95% confidence interval for ($\hat{\mu}_2 - \hat{\mu}_1$) is then given approximately by $(\hat{\mu}_2 - \hat{\mu}_1) \pm 1.96\sqrt{2} * s_p$. The null hypothesis is then rejected if $2ks_p > 1.96\sqrt{2} * s_p$ or $k > 1.4$, approximately. This result holds so long as the ratio of the standard errors does not vary appreciably, say by more than 2:1 (Goldstein and Healy, 1995; Goldstein, 1995).

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

Higher Education Funding Council for England (HEFCE) (1999a). Performance indicators in higher education: 1996-97, 1997-98: Report. Circular 99/66. Bristol: Author.

Higher Education Funding Council for England (HEFCE) (1999b). Performance indicators in higher education: 1996-97, 1997-98: Overview. Circular 99/67. Bristol: Author.

Kreft, I., and De Leeuw, J. (1998). Introducing Multilevel Modeling. Thousand Oaks, London, New Delhi: Sage.

Menard, S. (1995). Applied Logistic Regression Analysis. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-106. Thousand Oaks, CA: Sage.

O'Hagan, A. (1994). Kendall's Advanced Theory of Statistics, Vol. 2B: Bayesian Inference. London: Edward Arnold.

Ozga, J. and Sukhnandan, L. (1997). Undergraduate non-completion. Report No. 2 in Undergraduate Non-completion in Higher Education in England. Bristol: Higher Education Funding Council for England.

Pascarella, E.T., and Terenzini, P.T. (1991). How College Affects Students. San Francisco: Jossey-Bass.

Robinson, W.S. (1950). Ecological Correlations and the Behavior of Individuals. American Sociology Review, 15, 351-7.

Rodríguez, G. and Goldman, N. (1995). An assessment of estimation procedures for multilevel models with binary responses. Journal of the Royal Statistical Society, A, 158, 73-89.

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

Snijders, T. A. B. and Bosker, R. J. (1999). Multilevel Analysis: An introduction to basic and advanced multilevel modeling. London, Thousand Oaks, New Delhi: Sage.

Thomas, M., Adams, S. and Birchenough, A. (1996). Student withdrawal from higher education. Educational Management and Administration, 24, 207-21.

Tinto, V. (1987). Leaving College: Rethinking the Causes and Cures of Student Attrition. Chicago: University of Chicago Press.

Tinto, V. (1993). Leaving College: Rethinking the Causes and Cures of Student Attrition. (2nd ed.) Chicago: University of Chicago Press.

Tinto, V. (1995). Dropout from higher education: A theoretical synthesis of recent research. Review of Educational Research, 45, 89-125.

Tinto, V. (1997). Classrooms as Communities: Exploring the Educational Character of Student Persistence. Journal of Higher Education, 68, 6: 599-623.

Yorke, M. (1999). Leaving Early: Undergraduate Non-completion in Higher Education. London and Philadelphia: Falmer Press.

NOTICE

Reproduction Basis

- [x] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

- [ ] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").

EFF-089 (3/2000)