FOURTH INTERNATIONAL AMSTERDAM CONFERENCE ON MULTILEVEL ANALYSIS

Analysing student progress in higher education using cross-classified multilevel logistic models

John F Bell
Research and Evaluation Division
University of Cambridge Local Examination Syndicate
1 Hills Road
Cambridge
CB1 2EU
Tel. 01223 553849
Email: bell.j@ucles.org.uk

Abstract

Recently in the UK, there has been an increase in interest in the process of university admissions. Using archive data from the 1990s, the relationship between probability of obtaining a good degree and examination performance at age 18 will be considered. In addition, the issue of school type (state-maintained and independent) will also be considered. These archive data have been subject to various analyses before but the multilevel structure has been ignored. The data have a cross-classified structure with the students nested within universities and within schools. The models will be fitted using MCMC estimation techniques.

The paper will demonstrate the importance of exploratory analysis in formulating potential multilevel models. This paper will also consider the importance of school and university level variation.

Introduction

In recent years there has been a move towards more accountability in education. In the UK, one manifestation of this has been concerned with the issue of widening access to university, i.e., increasing the uptake of higher education by the lower social classes. In particular, concern has been expressed about the proportion of students who had attended independent (fee-paying) schools gaining access to the more prestigious universities. Some research that has been widely quoted suggests that students from independent schools tend to perform less well at university compared to students from state maintained schools (e.g., McNabb, Sarmistha and Sloane (1998) and Smith and Naylor (2001)). Unfortunately these papers ignore the multilevel structure of the data and so there is potential for the results to be misleading. The objective of this paper is to consider how multilevel modelling could be used to investigate issues relating to student progress.

To understand the issues it is necessary to consider a simplified structure of the higher education system as illustrated by Figure 1. At the bottom of the diagram are schools ordered in increasing quality from left to right. While at school, the students apply to universities in the belief that they will obtain the examination results required for the course they wish to study. In practice, this means universities set high entry requirements and can choose students, if the course is over subscribed and students can choose between courses with low entry requirements, if the courses are undersubscribed. However, there is a tendency for students from the better resourced schools (e.g. independent schools) to go on to study in the higher status universities on the top right hand side of the diagram. This results in a crossed multilevel structure and a hypothesis that students from the less well resourced state sector, obtain A-level results that do not reflect their true potential and so will not gain access to courses in the more prestigious universities.

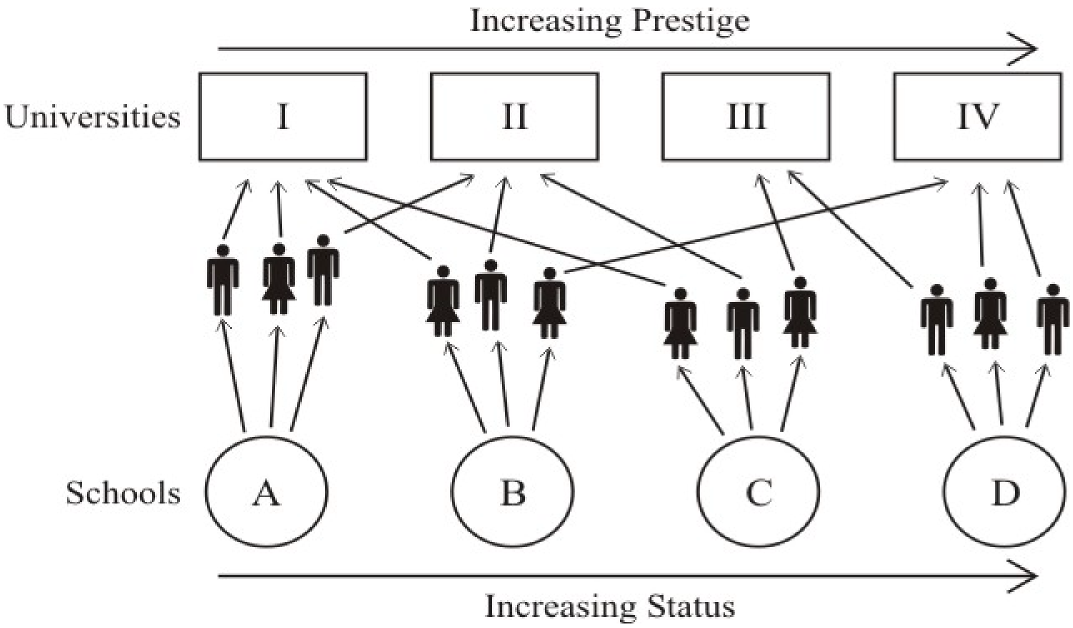

Figure 1: Simplified structure of Higher education system

Although the relative progress of students from different educational backgrounds is the source of much media interest, it is worth noting the age of the data used in this paper. The data comes from an archive of the University Statistical Record, which ended in 1993 as a result of a change to the higher education structure. In fact, there have been many changes in education since 1993. Now students from the state sector experience a school education that has been influenced by the introduction of the National Curriculum for ages 5-16 and its associated testing. The A-level (university entrance) examinations sat by these students are very different from the ones sat back then. The examinations are set by fewer awarding bodies (testing agencies) and are subject to greater government regulation. In addition, higher education has changed with an expansion of higher education and a reclassification of higher education institutions (this reform led to the change in the keeping of statistical records). Immediately after the last set of data was collected, institutions known as polytechnics were re-designated as universities. The students at these institutions would tend to have lower A-level grades than students in the 'old' universities that are considered here. The polytechnics are sometimes known as 'new' universities. This means that the results may not be generalizable to the current situation.

The university statistical record contains information for all students in all of the old universities in the United Kingdom. However, it was decided to use subsets of the data in this study. Obviously only students that had sat A-levels could be considered. It should be noted that the United Kingdom is made up of England, Scotland, Wales and Northern Ireland. The educational systems of these constituent parts vary. Although students from England, Wales and Northern Ireland attend Scottish institutions, both the Scottish school examination system and the university structure differ from those in the other parts. For this reason, it was decided that Scottish universities should be excluded from the analyses described in this paper. Students from Northern Ireland take A-levels but the organisation of the school system is very different from that of England and Wales and so the universities in Northern Ireland were also excluded.

The age of entry for the remaining universities is usually eighteen or nineteen. However, there are some students who start university when they are older. It is not unusual for these students to have relatively poor A-level examination results but this shortcoming has been compensated for by other qualifications and life experience in general. This means that they form a different population from the main body of the student population and the processes governing their selection and progress will be different. For this reason, a dummy variable was created to identify these students.

To measure prior attainment a score was derived from the A-level examination grades obtained by the candidates. Students take a wide variety of A-levels, usually of some relevance as a preparation for the course followed at university, although this varies from subject to subject. It would be surprising if the relationship between A-level score and university degree class was the same for all subjects. For this reason, it was decided to investigate the progress of students following the same course. Therefore the first step was that the most common courses were identified and from the list of the top ten, English, and mathematics were selected because they are both offered by a large number of universities, take students with a large range of prior attainment and are very different kinds of subject.

For the purposes of this paper, the A-level grades obtained have to be converted into a score. When the data used in this paper was gathered students intending to go to University usually sat three or four A-level examinations (the fourth examination usually being A-level general studies). These examinations are taken after a two-year course over the ages of 17 and 18. Traditionally pupils started at 11 and in what is known as the first form. The pupils taking A-levels are sometimes referred to as sixthformers (lower sixth formers for first year and upper sixth formers for the second and not sixth and seventh formers as might be expected). These examinations had five pass grades, A, B, C, D and E. For the purposes of this study, these A-level grades were converted into scores, 0, 2, 4, 6, and 10. (This is the usual tariff and was designed to allow for another examination that was deemed to be the equivalent of half an A-level examination). For each student, the total of the best three A-level grades excluding general studies was calculated. General studies was excluded because not all schools entered students for the examination and it was often ignored by universities. In this paper, this will be referred to as the A-level score. For the purposes of modelling, this score was standardised with mean 0 and variance 1 and will be referred to as the standardised A-level score.

Although there are a number of different types of schools that students attended to take A-levels, for the purpose of this paper, only the difference between state maintained and independent will be considered by using a dummy variable.

At the end of their courses students are awarded degrees which are classified as follows: Fail, ordinary, and four classes of honours degrees, 3rd, 2.2, 2.1, and 1st.
For the purposes of this paper, a binary dependent variable was formed taking the value 1 for a 1st or a 2.1 (sometimes referred to as a good degree) and 0 for all the other categories. This is consistent with some other analyses (e.g., Smith and Naylor, 2002).

In addition, one other variable was generated. Because the universities have different entry requirements it was decided that this needed to be considered in the modelling process. However because universities may make exceptions in special circumstances, the minimum A-level score is not a good indication of this requirement. It was decided to use the lower quartile of the A-level score for each institution as an approximation of the entry requirements.

In this paper, the last three years of data were considered (1991, 1992 and 1993) and dummy variables were created to allow investigation of year on year variation. Taking three years of data resulted in data sets of a reasonable size to demonstrate the analysis (for English, there were 33 universities, 673 schools and 6,654 students and for mathematics, there were 41 universities, 701 schools and 6,128 students).

The objective of this paper is to consider how data of this type could be analysed with multilevel models and not to provide a definitive analysis. There are more variables in the data set that could have been considered. After the data preparation described in this introduction, the next step is to carry out an exploratory analysis to guide the more formal modelling. This is described in the next section which is followed by a section on the fitting of multilevel models and finally the paper ends with a discussion of the implications of the results.

Exploratory Analysis

Before formally fitting multilevel models, an exploratory analysis was carried out with the package Arc (Bell, 2001, Cook and Weisberg, 1999). This is a highly interactive package and it is possible to consider large numbers of plots in a relatively short time. However, to illustrate this process only a small proportion of the plots considered will be presented. In the first, the relationship between the probability of obtained a good degree (goodeg) and the A-level score (best3ng). One of the graphs generated by the explanatory analysis is presented as Figure 2. In this graph two sets of lines has been plotted: one set for students from independent schools (the dotted lines) and another for students from the state sector. Each set consists of a logistic regression line base on a cubic fit and a LOWESS smoothed line. It is clear that in the subsequent multilevel modelling it is necessary to consider polynomial terms. The fact that the lines cross indicates that the relationship could be different for the two different student backgrounds.

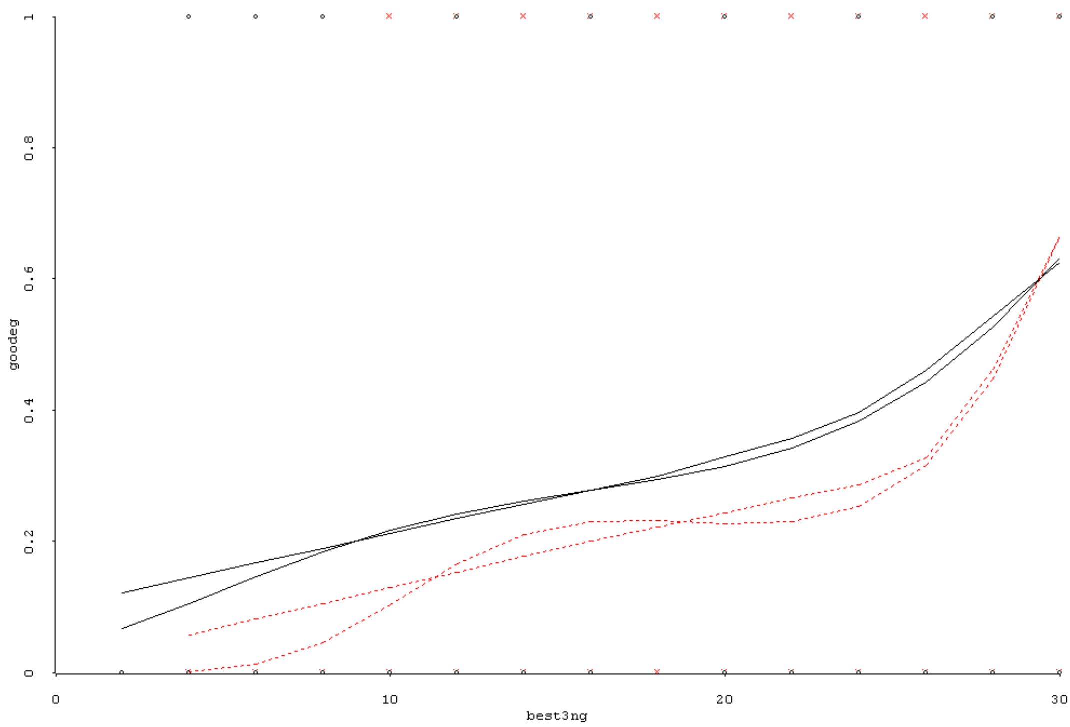

Figure 2: Exploratory data analysis of the relation between probability of good degree and A-level score by independent/state

In addition to the overall relationship, an inspection of the plots for individual universities (Figure 3(a)) showed that there was a considerable amount of variation between the universities and it would be necessary to consider complex models with random slopes. One feature of this plot is the fact that the lines sometimes show that candidates with very low A-level scores are very likely to obtain good degrees. The reason for this can be seen by considering the plots of each university. An example of these plots is given a Figure 3(b). In this figure, jittering (adding small amount of random variation to the coordinates) has been applied so that the individual data points do not overlap. This effect is caused by four candidates who were admitted with low A-level results but proved to be very able. These candidates had obviously been admitted on courses despite the fact they had not met the usual entry requirements. This group includes mature students who tend to do very well (not illustrated in the paper)

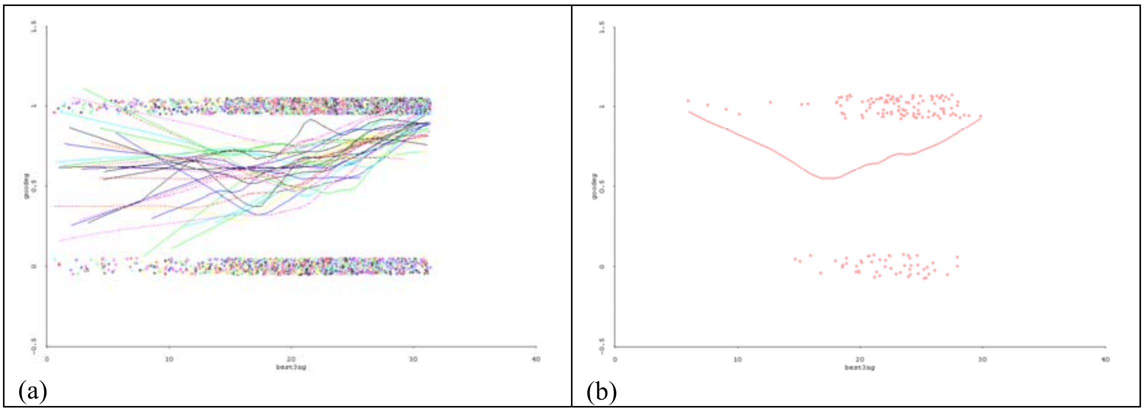

Figure 3: Lowess smooths for individual universities (English data set)

Differences between years were also considered. In Figure 4, the relationship between probability of obtaining a good degree and A-level score by year is considered for the mathematics data set. It is clear (if the figure is seen in colour), that 92 and 93 differ from 91.

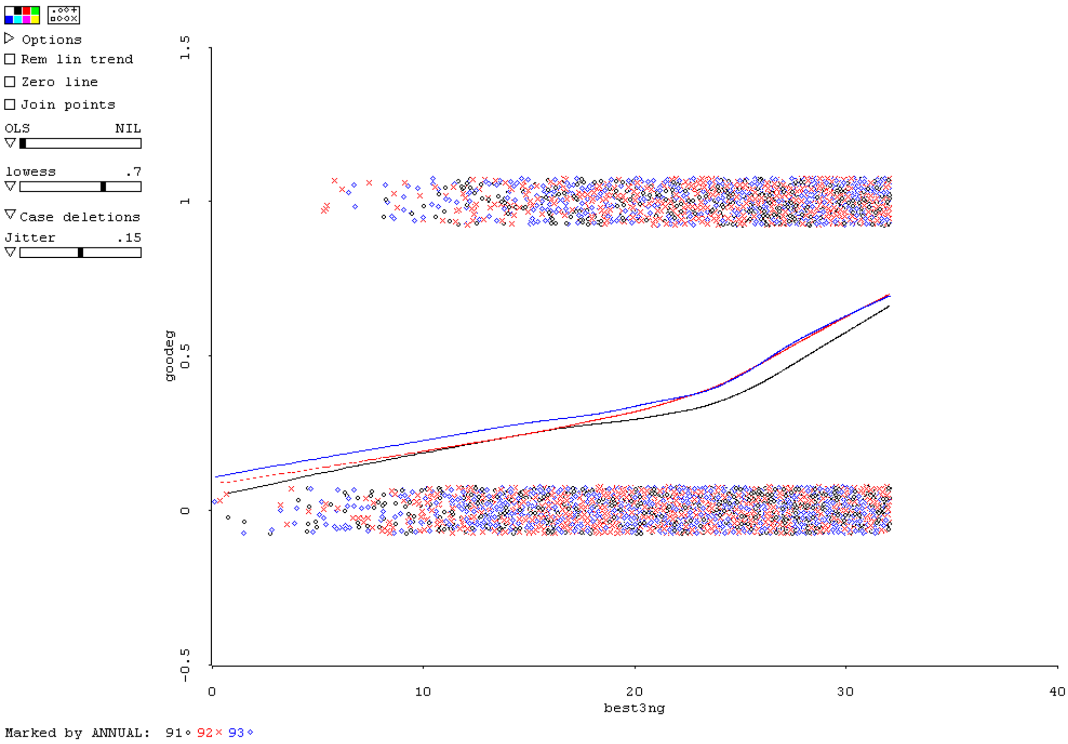

Figure 4: Lowess smooths by year (mathematics data set)

It is important to recognise that this type of exploratory analysis is only suggestive of what the features in the data are and some of the observed differences in the smoothed lines may not be statistically significant.

Results

To analyse the data sets compiled for this study, it is necessary to consider a multilevel cross-classified logistic regression model. These are complex models and the MCMC procedures in MLwin were used (it is the logistic aspect of the model that creates the problems, fitting cross-classified linear models is relatively straightforward with a sufficiently powerful computer). In addition to the cross-classified models, models ignoring university effects are also fitted to demonstrate the consequences of ignoring the true structure. Although many possible models were considered, only a small subset has been presented in this paper (for example, random slopes were considered but did not prove to be significant). The results of three analyses are presented in Table 1. For further information about the layout of the tables and the supporting figures see Bell (2002). In Model I only the school effects have been considered and the simple logistic model. Model II is the model with school level variation but the curves suggested by the exploratory data analysis. Finally model III is the analysis based on the cross-classified structure. The important difference between the first two models and the third can be seen in the random part. Ignoring the university level variation leads to the erroneous conclusion that schools have an influence on degree performance. The fact that the school level variation is zero for the cross-classified model suggests that the school level variation in models I and II is an artefact of the allocation of students to universities that differing relationship between degree success and A-level score.

Table 1: Parameter estimates for multilevel models of the English data set

Fixed Part

	Constant		Male (d)		Indep. (d)		Year 93 (d)		Mature (d)	
Model	Est.	s.e.	Est.	s.e.	Est.	s.e.	Est.	s.e.	Est.	s.e.
I	0.80	0.06	-0.17	0.06	-0.22	0.07	-		-	
II	0.40	0.07	-0.16	0.06	-0.29	0.08	0.13	0.06	0.82	0.12
III	0.41	0.10	-0.14	0.06	-0.31	0.08	0.15	0.06	0.72	0.12

Fixed Part (continued)

	A-level score		(A-level score)^2		(A-level score)^3		A-level * Indep	
Model	Est.	s.e.	Est.	s.e.	Est.	s.e.	Est.	s.e.
I	0.51	0.03	-		-		-	
II	0.50	0.07	0.23	0.03	0.05	0.02	0.19	0.08
III	0.42	0.08	0.24	0.04	0.06	0.02	0.21	0.08

Random part

	University		School		Student
Model	Est.	s.e.	Est.	s.e.	Fixed
I	-		0.06	0.03	1.00
II	-		0.06	0.03	1.00
III	0.16	0.05	0.00	0.00	1.00

Interpreting the coefficients of the fixed part of the model from the parameter estimates given in the table is not straightforward. In particular, the coefficient of a polynomial logistic regression with an interaction cannot be interpreted from a table and it is necessary to calculate the predicted probabilities and plot them. In Figure 5, the predicted probabilities for female candidates for the definitely incorrect model I and the possibly incorrect model II (the relationships could change with the addition of more variables) have been presented. This clearly indicates the need for checking the model assumptions when considering this type of model.

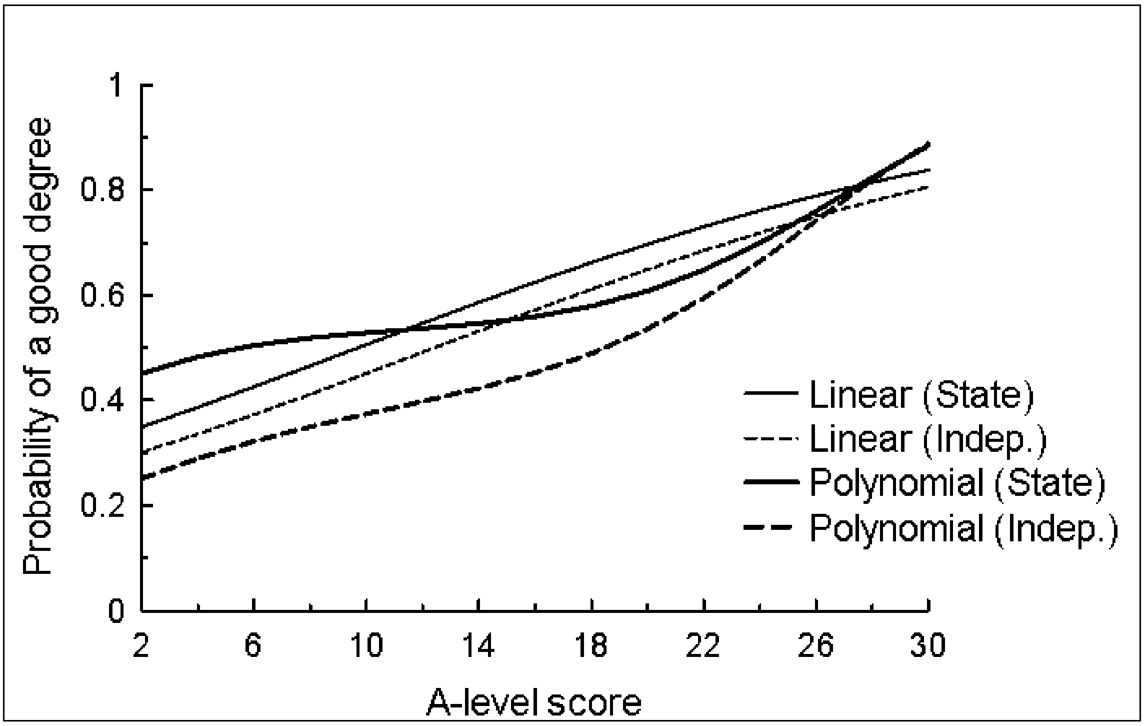

Figure 5: Comparison of model I and model III (English data set)

In Figure 6 the relationship between the sexes for model III has been presented. It is also useful to know the distributions of the different types of students. In this case, the greatest difference between students with different backgrounds occurs when the numbers of each are relatively small. The graph illustrates that at high levels of A-level performance there is no difference between the predicted probabilities for students from state and independent sector. Although the sex difference has been included in the figure, there are problems in interpreting it. This is the result of considering relative progress (i.e. the outcome variable being different from the measure of prior attainment). It is impossible to tell whether the difference is not the result of a sex difference in the A-level score or the assessment methods used in the degree or a difference in motivation or efficacy of study of the sexes. Similar arguments apply to the increased probability of success in 1993. The other significant parameter in the fixed effects indicates that mature students were more likely to obtain good degrees.

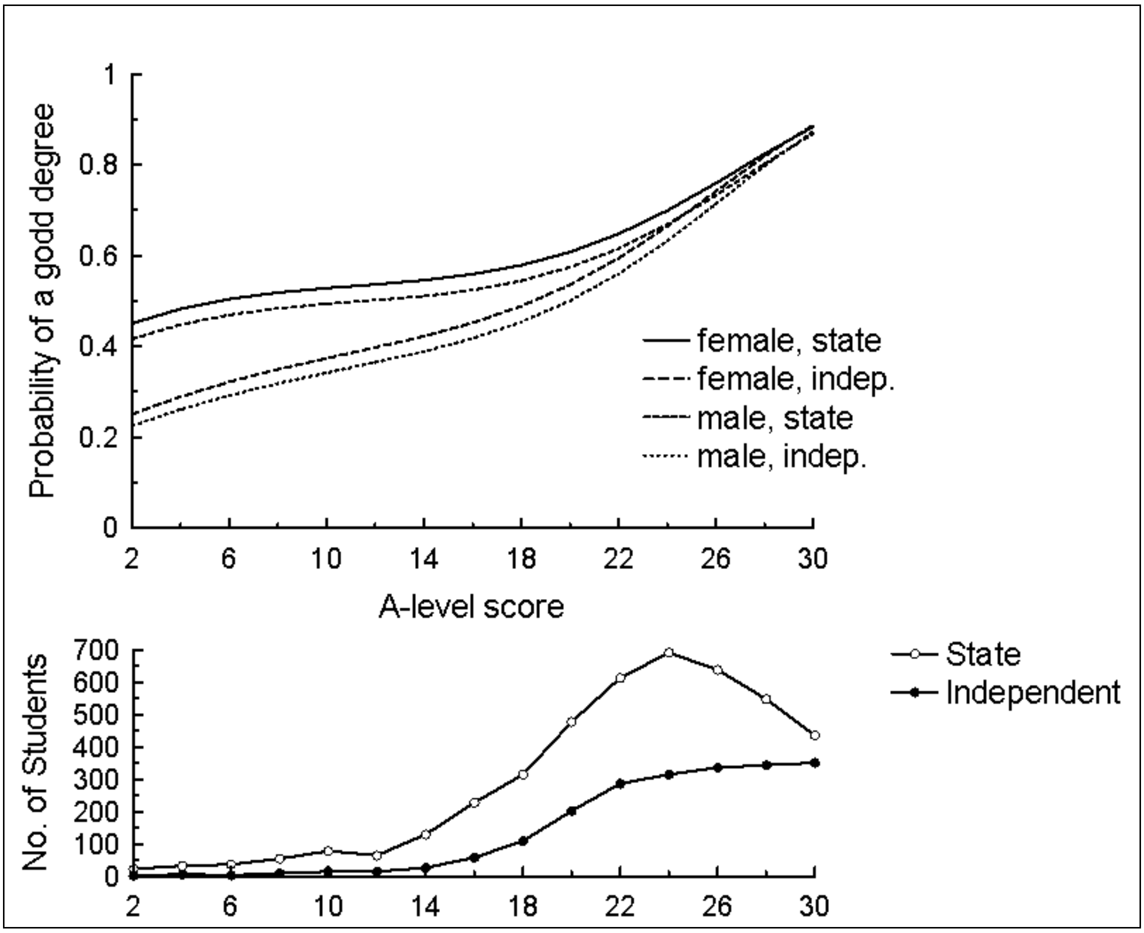

Figure 6: Predicted probabilities of obtaining a good degree for English

One feature of Figure 6 is of particular interest. Individual universities tend to have a student entry with the majority of students having A-level scores from a fairly restricted range. One consequence of this is that an English lecturer (assuming an interest in statistics) in an institution with the majority of applicants with an A-level score of 18 or less would conclude that A-levels were not a very good predictor of degree success. However, a lecturer at an institution with an entry with A-level grade scores of 18 or over could notice a stronger relationship.

There is also a significant amount of variation at the university level. This is presented in Figure 7. It is clear that the probability of obtaining a good degree varies considerably between universities and there is an obvious need for further research into the causes of these differences. Note it is a condition of access to the data set that this university is not identified without permission of the institution. Finally, there was no school level variation when the University level variation was included in the model. Apparent differences between schools occur because of a combination of the variability between universities and the difference in distribution of schools in individual university entries. This clearly demonstrates the need to consider the correct structure of the data set. Although the cross-classified model is not necessary, it was necessary to fit the model to determine that this was indeed the case.

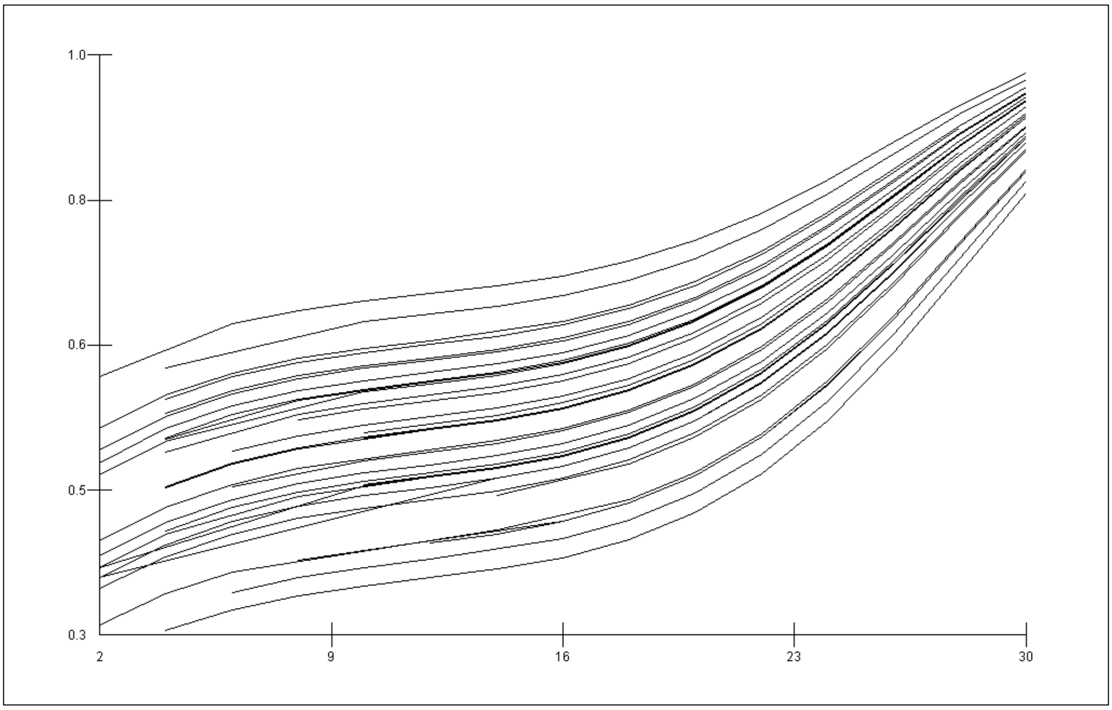

Figure 7: Variability at the University level for female state school students (95% Confidence Interval)

In Table 2, the parameter estimates for a subset of models fitted to the mathematics data set are presented for mathematics. The results for mathematics differ from those of English in a number of ways that are described below.

Table 2: Parameter estimates for models fitted to the mathematics data set

Fixed Part

	Cons		Sex (d:male)		Indep. (d)		Years 92,93 (d)		Lower quartile	
Model	Est.	s.e.	Est.	s.e.	Est.	s.e.	Est.	s.e.	Est.	s.e.
I	-0.79	0.05	0.20	0.06	0.15	0.07	-		-	
II	-0.30	0.14	0.21	0.06	-0.42	0.13	0.20	0.06	-0.08	0.01
III	-0.24	0.23	0.17	0.06	-0.47	0.14	0.20	0.06	-0.07	0.01
IV	0.01	0.18	0.22	0.06	-0.44	0.13	0.22	0.06	-0.06	0.01

Fixed Part continued

	A-level score		(A-level score)^2		(A-level score)^3		A-level*Indep.	
Model	Est.	s.e.	Est.	s.e.	Est.	s.e.	Est.	s.e.
I	0.70	0.03	-		-		-	
II	0.53	0.08	0.23	0.03	0.14	0.03	0.33	0.10
III	0.50	0.08	0.26	0.04	0.16	0.03	0.36	0.10
	0.48	0.09	0.26	0.04	0.17	0.03	0.26	0.10

Random Part

	University		School		Student
Model	Est.	s.e.	Est.	s.e.	Fixed
I	-		0.02	0.02	1.00
II	-		0.03	0.02	1.00
III	0.10	0.03	0.04	0.02	1.00
IV	0.06	0.02	0.01	0.01	1.00

It can be seen that for model III, there is both school and university level variation. However, an inspection of the university level residuals (Figure 8) reveals that there is one outlier – a university that would be predicted to have a large number of students with good degrees but did not. The students from this university were excluded and a series of analyses were carried out on the reduced data sets. This leads to Model IV which is discussed in detail below.

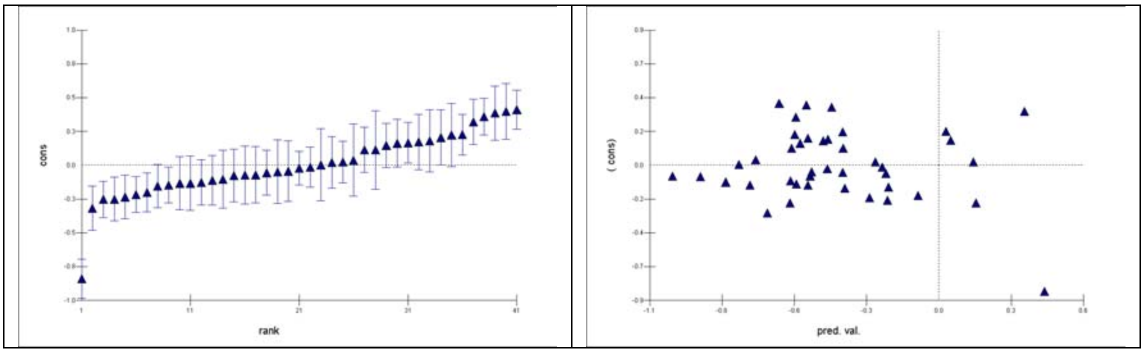

Figure 8: University level residual plots for mathematics model III

The relationship between the probability of obtaining a good degree and the A-level score is similar to that of English in that there is no difference for higher A-level scores, but there is a large difference in favour of the state schools for lower scores. The relationship is much stronger. For example for an A-level score of 12 the probability of obtaining a good degree is 0.43 for a male candidate from a state school and for a score of 26 it is 0.76. The equivalent figures for English are 0.50 and 0.73. However, there is a difference in the pattern of previous education. The number of independent school students taking mathematics is much lower than the number of state school students. In particular, students from the independent sector with poor grades tend not to study mathematics.

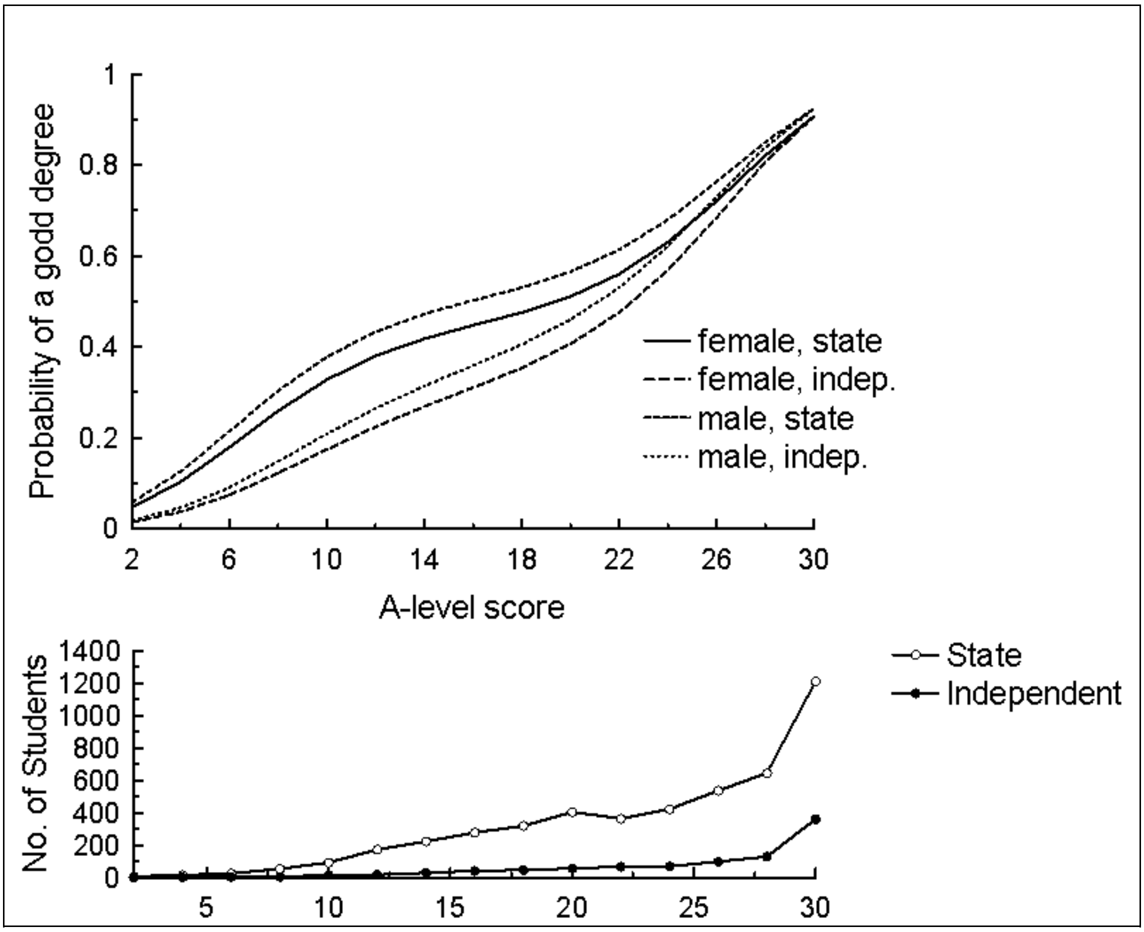

Figure 9: Relationship between probability of good degree and A-level score

There are also other differences between the models for English and Mathematics data sets. Being mature has no significant effect for mathematics. The gender effect has reversed. The 1991 cohort performed different from the 1992 and 1993 cohort. The most interesting effect is that the lower quartile was significant and inversely related to the probability of obtaining a degree. In general, a student with good A-level results would have a much higher probability of obtaining a good degree if she or he chose a course with lower entry requirements assuming that there was no unexplained university level variation. Of course, this is not the case. The relationship between the probability of obtaining a good degree and the lower quartiles has been present as Figure 10. Note that the fitted lines have been extended four points below the fitted lines and not all the fitted lines have been included in the plot.

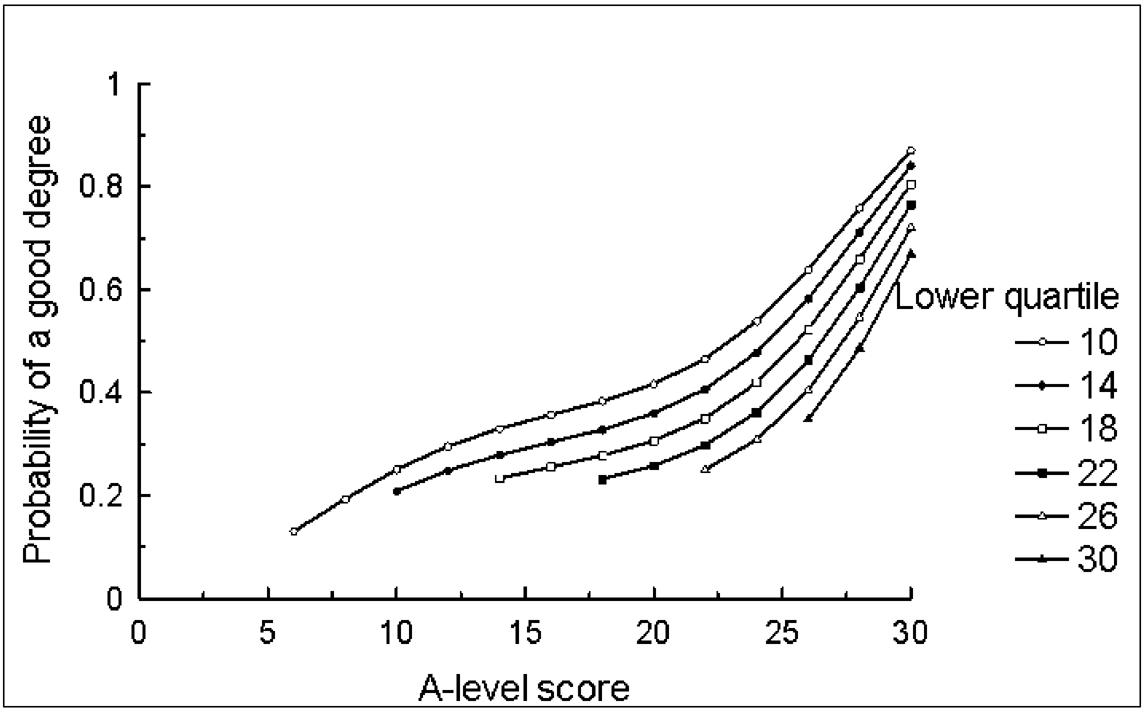

Figure 10: The relationship between the probability of obtaining a good degree and the lower quartile

Conclusion

Earlier in this paper it was noted that there had been many changes in the educational system since the data analysed in this paper were gathered. Therefore, the question that must be asked is what the effect has been of all the changes in education had over the last ten years? Do the results of an analysis of data from 1990-1993 still hold in 2003? If not, then this paper's only significance is that it demonstrates potential methods for analysing data on university progress. However, it is unlikely that the variation between universities has disappeared and any research that fails to consider the multilevel structure inherent in this type of data should be treated with suspicion. Also given the existence of interaction terms in this data, any analysis that does not state that such interactions have been considered should be treated with caution. There is also variation between subjects which should be considered. Also, it is possible that outliers can have an important effect of the fitted models so these needed to be considered. Any analysis of recent data that does not show evidence of these issues being considered should be considered with suspicion. In the United Kingdom, the university admissions process is coming under greater and greater levels of scrutiny and it is important that good quality statistical evidence is used to inform this debate.

If the patterns found in this data set still apply then there are a number of questions that need to be considered. Firstly, why is there such variability in the probability of obtaining a good degree? This question is directly related to the issue of comparability between awarding bodies for A-levels where it is recognised that there are difficulties in interpreting the statistical evidence (Bell and Dexter, 2000;

Bell and Dexter, 2000; Bell and Greatorex, 2000; Bramley, Bell, and Pollitt, 1998; Elliott and Greatorex, 2002; Elliott, Greatorex, Forster, and Bell, 2002; Forster and Gray, 2000. Jones, 1997).

From the methodological point of view it would seem that these data are very susceptible to model mispecification. This means that great care is needed in modelling the data and indeed further modelling might lead to results that differ in detail from those presented in this paper. However, one thing that is clear is that when sweeping generalisations are made about progress of students at university, they should be subjected to careful scrutiny. It is a good idea to remember Box's (1976) famous quote.

All models are wrong but some models are useful.

Acknowledgement

The data were supplied by ESRC data archive at the University of Essex. Any opinions expressed in this paper are solely those of the author and do not necessarily reflect those of the University of Cambridge Local Examination Syndicate or any of its subsidiaries.

References

Bell, J.F. (2001) *Visualising multilevel models: the initial analysis of data*. Third International Conference on Multilevel Analysis, Amsterdam, April.
Bell, J.F. (2002) *On the presentation of the results of multilevel analysis*. Society for Multivariate Analysis in the Behavioural Sciences, Tilburg, The Netherlands, July.
Bell, J.F. and Dexter, T. (2000) *Using Multilevel Models to Assess the Comparability of Examinations*. Fifth International Conference on Social Science Methodology of the Research Committee on Logic and Methodology (RC33) of the International Sociological Association (ISA), University of Cologne, October.
Bell, J.F. and Dexter, T. (2000) Using Ordinal Multilevel Models to Assess the Comparability of Examinations. *Multilevel modelling newsletter*, December, 12, 2, 4-9. (Available at http://multilevel.ioe.ac.uk/publref/new12-2.pdf).
Bell, J.F. and Greatorex, J. (2000) *A Review of Research into Levels, Profiles and Comparability.* QCA.
Box, G.E.P. (1976) Science and statistics. *Journal of the American Statistical Association,* 71, 791-799.
Bramley, T., Bell, J.F., and Pollitt, A. (1998) Assessing changes in standards over time using Thurstone Paired Comparisons. *Education Research and Perspectives*, 25, 2, 1-23.
Bratti, M. (2002). Does the choice of university matter? A study of the difference across UK universities in life science students' degree performance. *Economics of Education Review*, 21, 431-443.
Elliott, G. and Greatorex, J. (2002) A fair comparison? The evolution of methods of comparability in national assessment, *Educational Studies*, 28, 3, 253-264.
Elliott, G., Greatorex, J., Forster, M., and Bell, J.F. (2002) Back to the future: A methodology for comparing old A level and new AS standards. *Educational Studies 28,* 2, 163-180.
Forster, M. and Gray, E. (2000) *Impact of independent judges in comparability studies conducted by awarding bodies*. British Educational Research Association Annual Conference, Cardiff University, September.
Jones, B.E. (1997). Comparing Examination Standards: is a purely statistical approach adequate? *Assessment in Education*, 4(2), 249-262.
McNabb, R., Sarmistha, P., and Sloane, P. (1998*). Gender differences in student attainment: The case of university students in the UK*. mimeo, University of Cardiff.
Naylor, R., and Smith, J. (2002) *Schooling effects on subsequent university performance: Evidence for the UK university population*. Warwick Economic Research Papers. No. 657. Warwick: Department of Economics
Smith, J., and Naylor, R. A. (2001). Determinants of degree performance in UK universities: a statistical analysis of the 1993 student cohort. *Oxford Bulletin of Economics and Statistics*,63, 29-60.