Predicting Faculty Membership - Application of Student Choice Logit Model

Foula Kopanidis, School of Economics, Finance and Marketing, RMIT University, foula.kopanidis@rmit.edu.au

Abstract

This study investigates the application of a student choice logit model to examine the role demographic, socioeconomic and psychological variables play in influencing undergraduate students' preferences towards choosing to study in particular disciplines. As a tool of analysis, logit regression has offered an innovative approach that allows the researcher to estimate the probability that a prospective student will apply to certain disciplines. The likelihood ratio of the model is highly significant (p =.000) suggesting choice behaviour can be effectively explained by a set of particular explanatory variables. The marketing application of the logit model in generating 'student types' and simulation analyses is discussed. These findings have implications for higher education institutions wishing to develop segmented marketing strategies to influence choice behaviour of prospective students.

Introduction

The Australian undergraduate market is a large and profitable segment accounting for 66.5 percent of the total product and service offered by the Higher Education industry (IBIS 2010). In 2008-09 industry revenue amounted to $19.84 billion and this is expected to increase by 4.2 percent over the next five years. Every year school leavers represent almost 69 percent of students who make the decision to enter higher education (IBIS 2010). Accordingly, student cohorts applying for academic programs of their preference will consider different selection attributes when making their choice of program and university (Veloutsou, Lewis, and Paton, 2004). Student attrition rates impact on universities' enrolment levels, funding and policy formation (Krause, Hartley, James, and McInnis, 2005). Therefore, recruiting and retaining a financial interest in the share of the undergraduate market demands an understanding of the student characteristics that are significant in influencing student preferences by tertiary institutions (Long, Ferrier, and Heagney, 2006; Drewes & Michael, 2006; Maringe, 2006; James, Baldwin, & McInnis, 1999). This research seeks to contribute to understanding the drivers of prospective students' choices towards selecting to study in particular disciplines through the application of a logit model.

Review of the Literature

Prior literature on student choice behaviour suggests that the interactive effects of internal and external factors on a series of interrelated stages shaping students' choices can be integrated within multi-attributed models (Moogan and Baron; 2003; Harker, Slade, and Harker, 2001; Hossler and Gallagher, 1987). It appears that student's progress through at least three stages in their quest to realise their preferences for degree programmes and tertiary institutions (Chapman, 1986, 1981). Determinants of choice behaviour that impact on decision making, perceived risk and level of consumer

involvement consist of internal drivers, for example, values, motivation and selection criteria, (Dawes and Brown, 2002; Vallerand *et al* 1999; Homer and Kahle 1988; Rokeach 1971) and external drivers which are broadly captured by demographic and socioeconomic factors. Demographics are considered beneficial in predicting consumption behaviours (Kahle, Beatty and Homer, 1986), making them useful for segmentation purposes. McCarty and Shrum, (1993) have noted the importance of the inclusion of demographic factors (gender, age, income and education) in the explanation of behaviour.

Research Methodology

In an educational context, a number of discrete choice models (logit, nested logit, and probit) have been used to develop models of behavioural choice or of event classification. Logit models enable researchers to estimate the probability that a prospective student will express preferences towards particular choices. As a tool of analysis, logistic regression (multinominal logit in particular) has been used extensively for categorical dependent variables in the marketing literature for modeling consumer choice (McFadden 1974). Such a model allows the researcher to predict a discrete outcome such as in this research, group membership from a set of variables for each case, and to rank the relative importance of independents (Tabanachnick and Fidell, 2007).

The sample consisted of 304 first year undergraduate students enrolled in their first semester at an Australian university. A multi-stage cluster sampling approach was used to analyse each of the three faculties (Business n=128, Design and Social Context (DSC n=87) and Science and Engineering Technology (SET n=89), naturally representing a cluster in the population of interest. Specific degree programmes were then selected using a simple random probability technique using a sample of these clusters for surveying. A self administered survey of 21 questions constituted the measurement instrument drawn from previous surveys (Homer and Kahle 1988; Vallerand et al. 1992).

Student Choice Logit Model- Results

A multinomial logistic was employed to assess the prediction of membership in one of the three categories of outcome (Business, DSC and SET) first on the basis of demographic and socio economic predictors (baseline model) and second, with the addition of three metric drivers (full model). The likelihood ratio of the full model is highly significant (p =.000). Statistically significant relationships were identified amongst a number of independents and choice of faculty. These included the demographic variables of gender (p=.000), age (p=.000), country of birth, (p=.000), father's occupation, (p=.009), mother's education (p=.016), combined income (p=.002) and first preferences (p=.000). Psychological variables include external values (p= .042), motives of extrinsic (p=.000) and intrinsic (p=.003) and selection criteria of reputation (p=.000), entry (p=.026) and external (p=.020) influences. The pseudo coefficients of determination for the model are relatively high (Cox & Snell =.600, Nagelkerke .678 McFadde .423) Tabachnick and Fidell, 2007; Hernandez and Mazzon 2006). The model classification table indicates the student choice model predicting correctly almost 74% of the time, correctly classifying almost 81% for Business group,

84% for the DSC group and almost 54% for the SET group. In conclusion, the student choice model can be effectively used by marketers in education to predict the faulty type selected by prospective students on the basis of a set of explanatory variables.

Application of Student Choice Logit Model

To illustrate the marketing application of the student choice logit model, generated student types and a set of simulation analyses were conducted. Probability outputs for student classification were estimated on the basis of modal values for characteristics (refer to Table 1) and average values for derived factor variables (refer to Table 2).

Table 1. Modal Values for Demographic and Socio Economic Variables

Faculty	**Business**	**DSC**	**SET**
Gender	Female (1)	Female (1)	Male (2)
Age Aggregate	17- 20(1)	21- 27 (2)	17- 20 (1)
Country of Birth	Australia (1)	Australia(1)	Australia (1)
Preferences	Yes (1)	Yes (1)	Yes (1)
Occupation- Father	Professional (1)	Professional (1)	Professional (1)
Occupation- Mother	Working (4)	Professional (1)	Working (4)
Education – Father	University (3)	University (3)	University (3)
Education – Mother	At least High School (1)	University (3)	At least High School (1)
Combined Income	$80 000- $90 000+ (3)	$80 000- $90 000 +(3)	$50 000- $80 000 (3)

Table 2. Average Values of Derived Factor Variables

Faculty	**Business**	**DSC**	**SET**
Selection Criteria:			
Reputation Influence	0.2144719	-0.0440308	-0.2654126
Academic Influence	-0.0148854	0.0322597	-0.0101266
Entry Influence	0.0545619	0.1028874	-0.1790464
External Influence	0.3308722	-0.5446211	0.0565213
Motivation			
Extrinsic	0.2791487	0.5587757	0.1447467
Amotivation	0.2810578	-0.4772184	0.0622765
Intrinsic	-0.1908398	0.4350080	-0.1507664
Personal Values			
Internal	-0.0751254	0.4350080	0.0382182
External	0.2665618	-0.2653717	- 0.1239615

Consideration of a Typical Portfolio 'Student Type'

Based on computed average values, the student choice model was used to predict the probability of the three choice outcomes. Table 3 shows that for the typical Business student, a 99.99% accuracy rate was found; for a student enrolled in DSC, a 99.71% accuracy rate was predicted, and for a student enrolled in SET an accuracy rate of almost 75% was found. These prediction rates are proportional to the results specified from the classification matrix of the full model.

Table 3. Probability Output for the 'Typical' Faculty Student

Students Classification	Probability Values		
	Pr(BUS)	Pr(DSC)	Pr(SET)
Business	**0.999985**	1.38E-06	1.33497E-05
DSC	0.00174	**0.997176**	0.001084
SET	0.056757	0.194958	**0.748285**

To further investigate the effectiveness of the student choice logit model, two students enrolled in each faculty were randomly chosen and their particular profile constituted the basis for depicting a prediction change of this group membership. The model was particularly effective in predicting group membership for those students enrolled in both Business (92.6% and 99.7%) DSC (both 99%). However, as expected the choice model was an unimpressive predicting model for students enrolled in SET (48% and 38% respectively).

Application of Multinomial Logit Model: Simulation Analysis

To investigate what effects changes in derived factor scores have on the probability of predicting students' group membership, in this case the faculties of Business, DSC and SET, a set of simulation analyses were conducted (Oppenheim 1999). Using the estimated coefficients of predicator variables, a Microsoft Excel worksheet was constructed to compute the portfolio choice probabilities for student groups. For a given derived factor variable, the three outcome probabilities were computed over the range of values (-2 to + 2) for that factor variable. These outcome probabilities are computed setting all other variables in the model to their average or modal values. These computed probabilities are then plotted over the range of values for the factor variable. An example of a simulation change for the selection criteria of reputation is illustrated in figure 1. Clearly depicted is a change to the mean value of reputation influence from largely negative (*not at all important*) to strongly positive (*extremely important*) whereby the probability of being a classified as a student enrolled in DSC and SET steadily decreased, while the probability of being classified as a student enrolled in Business steadily increased. There is a clear cross over effect occurring when students enrolled in Business 'take over' students in both DSC and SET as the most likely students whom allocate particular importance to the selection criteria of 'reputation' influence.

Figure .1: Selection Criteria – Reputation Influence *(FA1REPINL)*

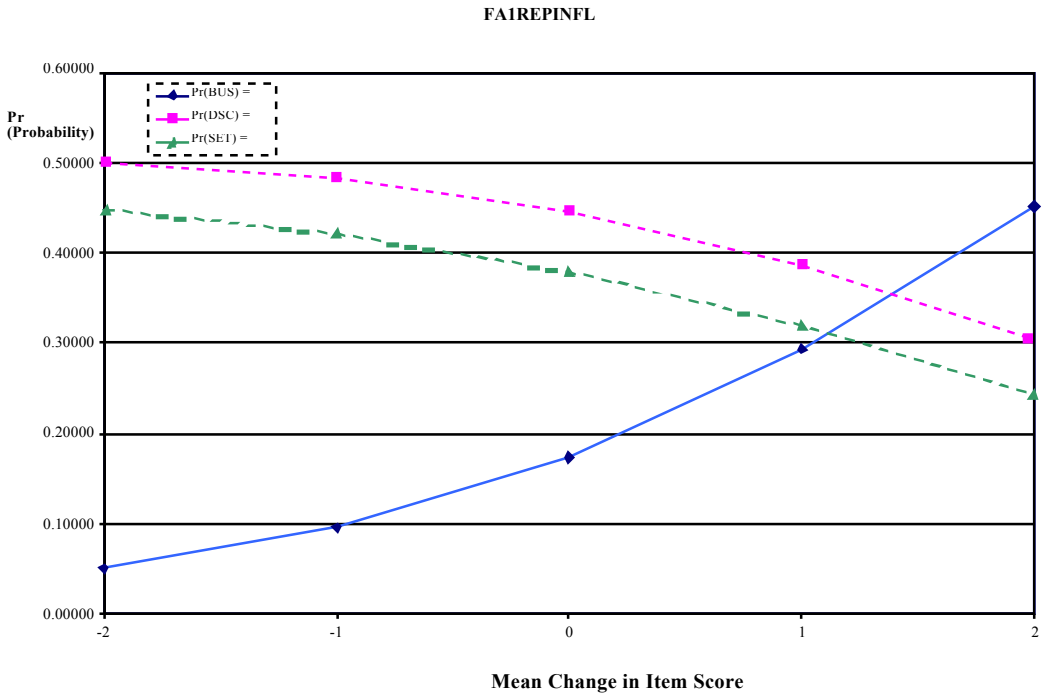

Table 4. Changes in rank order of cross over effects for FA1REPINFL

Ranking Values	Largely negative →	Largely positive
	1. DSC	1. Business
	2. SET	2. DSC
	3. Business	3. SET

Discussion and Conclusion

This study sought to adopt an innovative application of a student choice logit modelling through predicting the likelihood of student types preferring to enrol in a particular discipline. The results showed the constructed student choice logit model as effective in predicting correctly almost 74% of the time the chance of prospective new students expressing preferences to selecting a particular discipline to study in. The likelihood ratio of the full model is highly significant (p =.000) suggesting an undergraduate students choice of behaviour of selecting a particular discipline can be well explained by this a set of particular explanatory variables investigated in this study. Developing student types facilitates marketers to customise integrated marketing communication campaigns reinforcing the appeal to students who for example are strongly extrinsically or intrinsically driven. Appealing to relevant attributes of the program within disciplines increases the likelihood of attracting and retaining the 'right student type' most suited to the degree program offered.

Appendices

The student choice logit model is specified as

$$\pi_{ij} = \frac{\exp\left(X_i \beta_j\right)}{1 + \sum_j^J \exp\left(X_i \beta_j\right)},$$

where:

$\pi_i j$ is the probability that the *i*th student will choose the degree j^{th} program (Business as preferred degree program (y=1); DSC as preferred degree program (y=2); SET as preferred degree program (y=0))

X_i is a vector of independent variables (included a respondent's gender, age, country of birth, and preference selection of currently enrolled degree program and the socio economic influences of parental education, occupation and aggregate income).

References

Chapman, R., 1981. A model of Student College Choice. The Journal of Higher Education 52 (5) (Sep-Oct), 490-505.

Chapman, R., 1986. Towards a Theory of College selection; a model of college Choice search and choice behaviour. Advances in Consumer Research 13 (1), 246- 250.

Dawes, L.P., Brown, J., 2002. Determinants of Awareness, Consideration and Choice Set Size in University Choice. Journal of Marketing for Higher Education 12 (1), 49-73.

Drewes, T., Michael, C., 2006. How do students choose a university? An Analysis of Applications to Universities in Ontario Canada. Research in Higher Education 47(7), 781-800.

Harker, D., Slade, P., Harker, M., 2001. Exploring the Decision Process of School Leavers and Mature Students in University Choice. Journal of Marketing for Higher Education11 (2), 1-19

Hernandez, J.M., Mazzon, J.M., 2007. Adoption of internet banking: proposition and implementation of an integrated methodology approach. International Journal of Bank Marketing 25,(2), 72-88

Hossler, D., Gallagher, K. S., 1987. Studying student college choice: A three phase model and the implications for policymakers College and University 62 (3) 207-221

Homer, P., Kahle, L., 1988. A structural equation test of the 'Value-Attitude-Behaviour Hierarchy. Journal of Personality and Social Psychology 54 Winter; 638-664.

Kahle, L.R., Beatty S.E., Homer, P, 1986. Alternative Measurement Approaches to Consumer values: The List of Values (LOV) and Values and Lifestyle (VALS). Journal of Consumer Research 13, 405-409.

IBISWorld Industry Report (2010) Higher Education in Australia: N8431 February

James, R., Baldwin, G., McInnis C., 1999. Which University? The factors influencing the choices of prospective undergraduates, Centre for the Study of Higher Education, The University of Melbourne, Department of Education Training and Youth Affairs

Krause, K-L., Hartley, R., James, R. McInnis, C., 2005. *The First Year Experience in Australian Universities: Findings from a Decade of National Studies,* Centre for the Study of Higher Education University of Melbourne, Higher Education Innovation Programme, Department of Education, Science and Training

Long, M., Ferrier, F., Heagney, M. 2006. *Stay, play or give it away? Students continuing, changing or leaving university study in first year.* Centre for the Economics of Education and Training, Monash University ACER

Maringe, F., 2006. University and course choice Implications for positioning, recruitment and marketing. International Journal of Educational Management 20(6), 466-479.

Moogan, Y.J., Baron, S., 2003. An Analysis of Student Characteristics within the Student Decision Making Process', Journal of Further and Higher Education, 27(3), 271-287.

McCarty, J.A., Shrum, L.J., 1993, The Role of Personal Values and Demographics in Predicting Television Viewing Behaviour : Implications for Theory and Application', Journal of Advertising, XXII (4), 78-101.

McFadden, D., 1974. Conditional logit analysis of quantitative choice behaviour', in Zarembka, P. (Ed.), Frontiers of Econometrics, Academic Press, New York, NY.

Rokeach, M., 1973). The nature of human values, The Free Press, New York, NY

Soutar, G., Turner, J., 2002. Student's preference for university: a conjoint analysis. The International Journal of Educational Management 16 (1), 40-45.

Tabanachnick, B.G., Fidell, L.S, 2007. Using Multivariate Statistics Allyn and Bacon, Boston MA.
Vallerand, R.J., Pelletier, L.G., Blais, M.M., Nathalie, N, M., Senecal, C., Vallieres, E.F., 1999. The Academic Motivation Scale: A Measure of Intrinsic, Extrinsic and Amotivation in Education. Educational and Psychological Measurement 52, 1003-1017.
Vallerand, R.J. Pelletier, L.G. Blais, M.M. Nathalie, N, M. Senecal, C. and Vallieres, E.F The Academic Motivation Scale: A Measure of Intrinsic, Extrinsic and Amotivation in Education', *Educational and Psychological Measurement* Vol 52 pp. 1003-1017
Veloutsou, C., Lewis J., Paton, R., 2004. University selection: information requirements and importance. The International Journal of Educational Management 18 (3), 160-171.