

USING MULTINOMIAL LOGISTIC REGRESSION MODEL FOR THE STUDYING MAJOSELECTION OF STUDENTS

Nguyen Nhu Trang

TNU –Thai Nguyen University of Economics & Business Administration

Dang Phi Truong

TNU –Thai Nguyen University of Economics & Business Administration

1. Abstract

This paper applies Multinomial regression logit model to identify factors influencing students' decision to choose a major. The study is conducted at Thai Nguyen University of economics and business administration in Vietnam, with 344 observations of 3rd and 4th year students. In which, the faculties are divided into 3 groups according to majors: (i) accounting, (ii) business, (iii) economics. The results showed that the accounting majors are often the choice of female students and have a hard-working attitude, while the factors of learning environment and curriculum information influence the decision of students majoring in business administration and economics.

Keywords: *Multinomial regression logit, students 'selection, accounting, business, economics...*

2. Introduction

Today, many universities are facing a huge decline in student populations. The competition is fierce not only between vocational schools but also among faculty majors in a school. The financial proactive mechanism is the reason why many schools and faculties in universities have to improve their competitiveness and attract learners. At the same time, the trend of choosing a career of students is a signal for schools to prepare or change the structure of a major. Most of previous studies which only research on the student's university choice decision, while a studying major choice of student in a school are rarely interested in research. Therefore, this paper is designed to consider the factors that influence the students 'selection of a major. The found out will cash light on for educational managers have basic to develop appropriate training strategies which meet the needs of learners and attract students.

3. Data and methodology

3.1 Data

The research data is collected from March 2020 to April 2020 by three channels: academic advisors (homeroom teachers), student volunteers, and the Student Union. The result, 357 responses from online have collected. After checking, 344 observations which eligible for analysis were retained. The survey questionnaire is built in 2 parts, part 1 is the demographic information; part 2 includes questions related to factors affecting the student's choice of studying major. The scale used is a 5-level Likert scale: level 1 is completely disagreeing, level 5 is completely agreeing.

3.2. Methodology

Multinomial logistic regression. Choice behavior is an issue that is of concern to many researchers. The student's major choice is multidisciplinary, so the dependent variable is the multiple choice variable. Greene (2012, 763–766), Hosmer, Lemeshow, and Sturdivant (2013, 269–289), Long (1997, chap. 6), Long and Freese (2014, chap. 8), and Treiman (2009, 336–341) introduced the Multinomial logit model. This article applies Multinomial logistic regression to identify factors affecting students' choice of majors at Thai Nguyen University of Economics and Business Administration.

The Multinomial logistic regression model has been used to estimate the significant factors that determine the probability of students' choice of studying major. Assuming that each student has a sole alternative among major choice options, and their choices are affected by factor X. The Multinomial logistic regression model for studying major choice specifies the following relationship between the probability of opting for Y_i and the set of explanatory variables X as follows (Greene, 2011):

$$Prob(y_i = j) = \frac{e^{\beta_j^1 X_i}}{\sum_{j=0}^3 e^{\beta_j^1 X_i}}$$

Where Y is the latent variable on the observed choice of alternative j by the ith individual, j ranges between 1 and 3, with 1 “accounting,” 2 “business,” 3 “economics”; P_{ij} is the probability that the individual chooses alternative j; X_i is independent variables (gender, place, tuition policy, learning environment, providing information, influence of relatives, ...); and β_j is a vector of coefficients on each X.

To interpret the effects of explanatory variables on the probabilities, marginal effects are usually derived as (Greene, 2011):

$$\delta_j = \frac{\partial P_j}{\partial i} = P_j \left[\beta_j - \sum_{k=0}^j P_k \beta_k \right] = P_j (\beta_j - \bar{\beta})$$

The marginal effects measure the expected change in probability of a particular choice being made with respect to a unit change in an explanatory variable (Greene, 2011). The signs of the 344 studying major choice of students marginal effects and respective coefficients may be different, as the former depends on the sign and magnitude of all other coefficients. The model is estimated using the maximum likelihood method. Additionally, since the estimated coefficients in a non-linear model are not interpreted directly, this paper relies on the marginal effects to better understand the magnitude of the coefficients (Greene, 2011). The estimation is carried out by STATA 14.1.

Factor analysis. For multi-item constructs (i.e. students’ evaluation), factor analysis is one of the most widely used approaches to reduce the number of variables to eventually produce a small number of principal components that could be ready for evaluation. Prior to performing factor analysis, Cronbach’s α is used to estimate the reliability or consistency of the scale.

3.3. Variable and scale

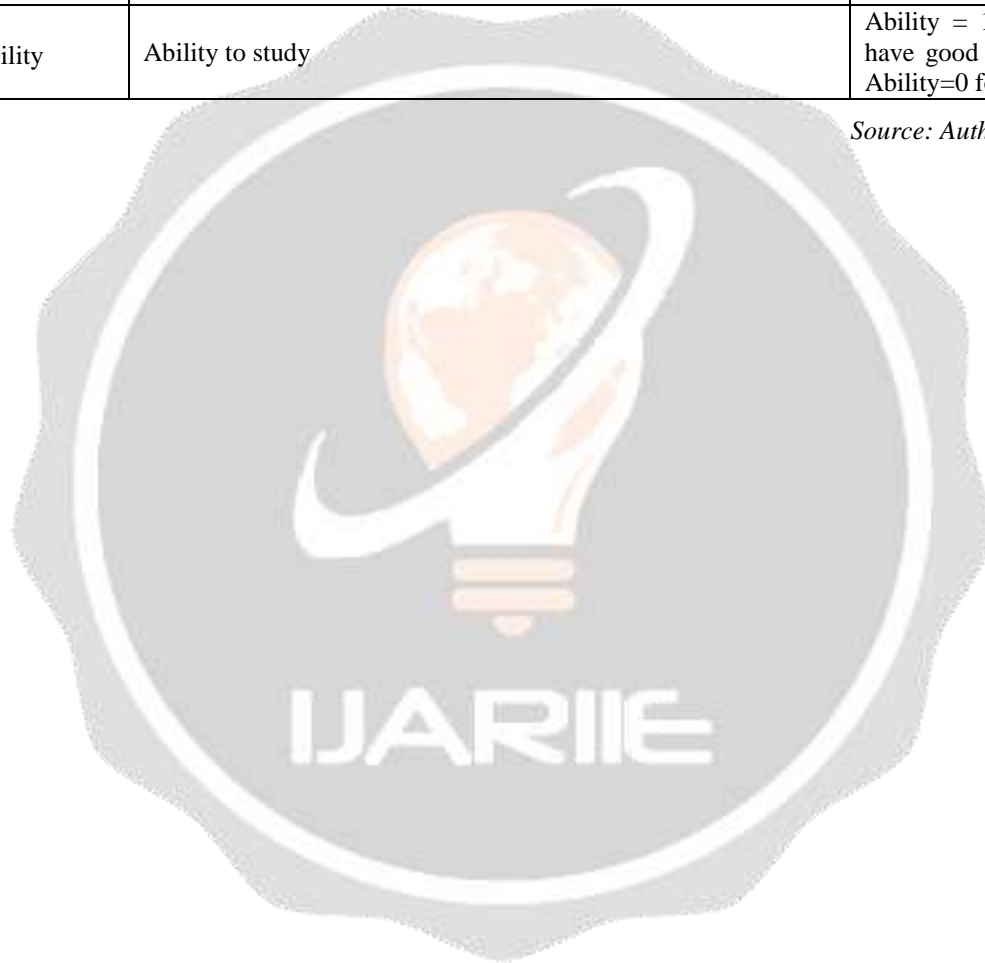
7 groups of factors with 31 items are selected and proposed in this study. These factors are described in Table 1.

Table 1. Variable and scale

Variable name	Describe variable	Scale
Tuition policy (CS)	(CS1) Reasonable tuition policy. (CS2) Reasonable living and studying expenses. (CS3) There is a policy of financial support (scholarships, grants, concessional loans...). (CS4) Flexible fee collection regime.	Likert 5
Curriculum (CT)	(CT1) Courses/subjects with content and structure and variety for students to choose from. (CT2) Have flexible entry registration procedures. (CT3) Advanced/intensive study programs tailored to students' needs. (CT4) Study program with many practical contents to meet the needs of students. (CT5) Allow flexibility when switching disciplines. (CT6) There are many training systems. (CT7) There are subjects/programs of study available for students to choose from and study throughout the course.	Likert 5
Reputation (DT)	(DT1) The faculty has a reputation for training. (DT2) The faculty has quality and prestigious programs. (DT3) The educational programs are recognized/appreciated for the quality of their training.	Likert 5
Learning Environment (MT)	(MT1) Training environment that encourages learning for students. (MT2) Facilities and equipment to serve students well for relaxation and entertainment. (MT3) Extracurricular activities, social activities, soft skill development, association activities, groups ... are diverse and abundant. (MT4) Resources needed to well-meet the learning needs of students. (MT5) Safe and clean learning environment. (MT6) Highly qualified and experienced staff and lecturers	Likert 5
Providing Information (TT)	(TT1) The faculty provides full information related to career opportunities. (TT2) The faculty provides a full range of information related to the courses. (TT3) The faculty provides full information related to graduate- level to study at a higher level.	Likert 5

Influence of relatives (AH)	(AH1) Influence from the advice of parents. (AH2) Influence from my friends' advice. (AH3) the Influence from the advice of students and alumni. (AH4) Influence from the advice of a high school teacher. (AH5) Influence from the advice of an admission counsellor.	Likert 5
Major	The studying major choice of students	Major = 1 if students choose accounting, Major = 2 if students choose business, Major = 3 if students choose economics
Gender	The gender of student	Male = 1, female =0
Place	The place students live	Place = 1 if students live in Thai Nguyen, Place = 0 otherwise
Ability	Ability to study	Ability = 1 if students have good study result, Ability=0 for otherwise

Source: Authors synthesiz



4. Research result

Before taking the Multinomial logistic regression for data, we used EFA for reliability for items of factors which will be used in the students' choice model. The found out, Cronbach's Alpha coefficient of 07 main variables is bigger than 0.6, so the selected items are reliable to perform EFA analysis. KMO value is 0.97 (>0.5), Barlett's test result is 9605,384 at significance level Sig = 0.000 <0.05, indicates that EFA analysis is consistent with research data. After rotated matrix was taken, 27 items which belonged to independent variables have divided into 3 groups of factors (MT, CS, AH). Where MT includes CT, MT, DT, TT. The total variance cumulative of 72.53% mean these 3 factors explain 72.53% of the variation of the data, Thus, these factors used to explain the scale for the selection decision of students are reasonable.

The selection of students in studying major are multi-choice, so the results was shown by comparing three selections. In this research, we estimated the factor can affect to students' selection in compared among accounting, business and the remain majors. The outcomes are followed in Table 2, Table 3, Table 4.

Table 2. The factor affected to the student's choice, base outcome of economic major

Major	Coef.	Std. Err.	z	P>z	[95% Conf. Interval]	
Accounting						
Gender	-2.624438	0.4932	-5.32	0	-3.5911	-1.65778
Place	0.4354772	0.3618	1.2	0.229	-0.27383	1.144788
Ability	1.079761	0.3954	2.73	0.006	0.30469	1.854832
Learning environment	-0.5531222	0.2524	-2.19	0.028	-1.04794	-0.0583
Providing information	0.380225	0.2589	1.47	0.142	-0.12733	0.887784
Influence of relatives	0.0605915	0.1678	0.36	0.718	-0.2684	0.389586
_cons	0.3530552	0.3262	1.08	0.279	-0.2864	0.992513
Business						
Gender	-0.4442823	0.2997	-1.48	0.138	-1.03174	0.143173
Place	0.2358657	0.3162	0.75	0.456	-0.38393	0.855662
Ability	0.5429047	0.36904	1.47	0.141	-0.18041	1.266218
Learning environment	-0.6021426	0.2247	-2.68	0.007	-1.04273	-0.16156
Providing information	0.4623718	0.2337	1.98	0.048	0.00424	0.920503
Influence of relatives	-0.09428	0.1501	-0.63	0.53	-0.38858	0.200024
_cons	0.697096	0.2917	2.39	0.017	0.125294	1.268898
Economics	(base outcome)					

Source: Authors calculated

In compared with Business (Major = 2) and Economics (Major =3), the female students and good studying result students are likely to choose accounting more than other. This reflects the fact that students who choose accounting major are often female, who study hard and often have good academic results. While the male students and lower studying results often participate in business and economics.

In compared with economic students, the providing information factor has positive effect on business student's selection and the learning environment factor has negative effect (Table 2, Table 4).

Table 3. The factor affected to the student's choice, base outcome of business major

Major	Coef.	Std. Err.	z	P>z	[95% Conf. Interval]
Accounting					

Gender	-2.1801	0.4573	-4.77	0	-3.07664	-1.28367
Place	0.1996	0.2973	0.67	0.502	-0.38318	0.782401
Ability	0.5368	0.2848	1.88	0.059	-0.02142	1.095135
Learning environment	0.0490	0.2070	0.24	0.813	-0.35677	0.454807
Providing information	-0.0821	0.2162	-0.38	0.704	-0.50604	0.341746
Influence of relatives	0.1548	0.14186	1.09	0.275	-0.12319	0.432928
_cons	-0.3440	0.2716	-1.27	0.205	-0.87646	0.188376
Business	(base outcome)					
Economics						
Gender	0.4442	0.2997	1.48	0.138	-0.14317	1.031737
Place	-0.2358	0.3162	-0.75	0.456	-0.85566	0.383931
Ability	-0.5429	0.3690	-1.47	0.141	-1.26622	0.180409
Learning environment	0.6021	0.2247	2.68	0.007	0.161558	1.042727
Providing information	0.4623	0.2337	1.98	0.048	0.9205	0.00424
Influence of relatives	0.0942	0.1501	0.63	0.53	-0.20002	0.388584
_cons	-0.6970	0.2917	-2.39	0.017	-1.2689	-0.12529

Source: Authors calculated

Table 4. The factor affected to the student's choice, base outcome of accounting major

Major	Coef.	Std. Err.	z	P>z	[95% Conf.Interval]	
Accounting	(base outcome)					
Business						
Gender	2.180155	0.4573	4.77	0	1.283669	3.076642
Place	-0.1996115	0.2973	-0.67	0.502	-0.7824	0.383178
Ability	-0.5368564	0.2848	-1.88	0.059	-1.09514	0.021423
Learning environment	-0.0490204	0.2070	-0.24	0.813	-0.45481	0.356767
Providing information	0.0821468	0.2162	0.38	0.704	-0.34175	0.50604
Influence of relatives	-0.1548715	0.1418	-1.09	0.275	-0.43293	0.123185
_cons	0.3440408	0.2716	1.27	0.205	-0.18838	0.876458
Economics						
Gender	2.624438	0.4932	5.32	0	1.657778	3.591097
Place	-0.4354772	0.3618	-1.2	0.229	-1.14479	0.273834
Ability	-1.079761	0.3954	-2.73	0.006	-1.85483	-0.30469
Learning environment	0.5531222	0.2524	2.19	0.028	0.058302	1.047942
Providing information	0.380225	0.2589	1.47	0.142	0.88778	0.127334
Influence of relatives	-0.0605915	0.1678	-0.36	0.718	-0.38959	0.268403
_cons	-0.3530552	0.3262	-1.08	0.279	-0.99251	0.286402

Source: Authors calculated

To estimate the effect of factor to the student's choice, the average marginal effect is conducted in this paper. The results shown in Table 5 have pointed out that the factors can affect to student are gender, ability, learning environment, providing information.

Table 5. The marginal effect of factors affected to the studying major choice of student

Major	dy/dx	Std.Err	z	p>(z)	[95% conf. Interval]	
Accounting						
Gender	-0.4605	0.0773	-5.95	0.000	-0.6122	-0.3088
Place	0.0534	0.0564	0.95	0.344	-0.0571	0.1640
Ability	0.1384	0.0544	2.54	0.011	0.0316	0.2452
Learning environment	-0.0250	0.0389	-0.64	0.521	-0.1013	0.0513
Providing information	0.0103	0.0404	0.26	0.799	-0.0690	0.0896
Influence of relatives	0.0254	0.0266	0.96	0.339	-0.0267	0.0776
Business						
Gender	0.2608	0.0702	3.71	0	0.1230	0.3986
Place	-0.0027	0.0617	-0.05	0.964	-0.1238	0.1182
Ability	-0.0173	0.0638	-0.27	0.786	-0.1425	0.1078
Learning environment	-0.0720	0.0434	-1.66	0.098	-0.1572	0.0132
Providing information	0.0616	0.0458	1.35	0.179	-0.0281	0.1514
Influence of relatives	-0.0321	0.0298	-1.08	0.281	-0.0905	0.0263
Economics						
Gender	0.1996	0.0491	4.06	0	0.1033	0.2959
Place	-0.0506	0.0498	-1.02	0.309	-0.1483	0.0469
Ability	-0.1211	0.0573	-2.11	0.035	-0.2335	-0.0086
Learning environment	0.0970	0.0348	2.79	0.005	0.0288	0.1653
Providing information	0.0719	0.0362	1.98	0.047	0.1430	0.0008
Influence of relatives	0.0066	0.0234	0.29	0.775	-0.0392	0.0525

Source: Authors calculated

Gender, increasing one more male student, the probability of choosing accounting major will reduce 46% choosing of business and economics will be increased by 19,96% and 26,08% respectively, at a statistic significant level of 99%.

Ability. The more students have good studying results, the more likely they want to choose the accounting major. At significant level 95%, increase one student has good results increase 13,84% accounting students and reduce 12,1% and 1,7% business and economic students at the same time.

The learning environment has a positive reaction to the business student's choice with 9,7% (at significant level of 99%), while this factor has negative to the selection of economics students. This can point that maybe the learning environment of the Faculty of Economics is not really favorable or comfortable for learners.

Finally, seemingly the economic faculty has provided information better than two remains faculty, or economic students are often more interested in providing course information. This is reflected by the positive impact of the Providing information variable on the student's choice, when the information provided is higher than the average level, the decision to choose to study economics major will increase to 7.2%

5. Conclusion

The results of factors influencing students' choice of majors found from the use of the Multinomial logistic regression model have suggested the enrollment strategy of Tueba faculties. The research results show that the accounting faculty should focus on attracting female students, while the faculties of business and economics need to improve their learning environment and provide more complete information to learners to increase the number of learners in these two faculties.

REFERENCES

1. Greene, W.H. (2011), *Econometric Analysis*, 7th ed., Pearson Prentice-Hall, NJ.
2. Greene, W. H. (2012), *Econometric Analysis*. 7th ed. Upper Saddle River, NJ: Prentice Hall.
3. Hosmer, D. W., Jr., S. A. Lemeshow, and R. X. Sturdivant (2013), *Applied Logistic Regression*. 3rd ed. Hoboken, NJ: Wiley
4. Long, J. S. (1997), *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage.
5. Long, J. S., and J. Freese (2014), *Regression Models for Categorical Dependent Variables Using Stata*, 3rd ed. College Station, TX: Stata Press.