

Effects of Demographic Variables and Academic Performance on Student Retention

Fumei Weng^{1, a *}

¹ Dept. of Applied Digital Media, WuFeng University, 117, Sec 2, Chiankuo Rd, Minhsiung, Chiayi County 62153, Taiwan

^afweng@wfu.edu.tw

* Corresponding author: Fumei Weng

Keywords: academic performance, first-year persistence; student retention; logistic regression.

Abstract. A model of the retention of first-year higher education students in Taiwan was created using logistic regression. The data used for creating the model consisted of the records of 955 first-year students who were enrolled in the academic year 2013 at a private institution in Taiwan. Six variables were found to be significant for determining student retention, namely: major field of study, age, residence, first-semester credit, second-semester grade and absenteeism. Seven logistic regression models were created using the six predictors obtained and the model with all the six predictors showed the best predictive power. Second-semester grade, residence and first-semester credits were all found to have a significant impact on first-year student retention.

Introduction

The disproportionate number of students who leave higher education between the first and second years is a major focus for retention studies. Research has also emphasised the fact that the first year of studies is a critical time for students to drop out [1-3]. In Australia, a study found that one third of all students failed to graduate and about half of these dropout students withdrew in their first year [4]. Student performance in the first year of studies is the strongest retention predictor for new students in their first semester [5, 6]. Thus, first-year student retention has become a major concern in higher education.

In Taiwanese higher education institutions, it is more viable to place more effort on retention activities than recruitment. There was also a big revolution in the Taiwanese education system in 2000 in the form of multiple entrance channels to universities. Multiple entrance channels increased the diversity of the student population in Taiwan and this shift in demographics in higher education should affect how higher education policy makers view student retention in the future. In order to understand why students withdraw from their studies, it is important to examine the diversity of students and their various characteristics. To predict the retention of the new generation of students in Taiwan, several attributes of first-year students were collected at the end of their first year of studies and were used to examine the persistence of the increasingly diversified student population.

Methodology

Data. The dataset used in the study contained the records of 955 first-year students who were enrolled in the 2013 academic year, which were provided by a private institution in Taiwan. The dataset included fourteen independent variables and one dependent variable as described in Table 1.

Method. In this study, outcome of study (OUTCOME) was the dependent dichotomous variable which represented either the dropout (OUTCOME=0) or persistence (OUTCOME=1) of the student in his/her studies. Statistical tests, namely cross-tabulation and correlation, were performed to test the relationships between the fourteen independent variables and the dependent variable. Logistic regression was used to predict student retention. In order to examine the effects of the predictors on student retention, the model including all predictors was used as the control model. Other models

were created by excluding one predictor at a time. The dataset was randomly partitioned into two parts: 70 per cent for modelling and 30 per cent for testing prediction accuracy.

Table 1. Student Variables

Variables	Values and Description
MAJOR	1 = Engineering, 2 = Business, 3 = Social Science, 4 = Security Science
SEX	1 = male, 2 = female
AGE	1 = < 18 yrs, 2 = [19, 21], 3 = > 22 yrs.
Secondary school (SSCHOOL)	1 = high school, 2 = vocational high school
Entrance test score (ESCORE)	[0.0, 100.0]
Special admission status (SASTAT)	1 = general admission, 2 = special admission
Residence (RESIDENCE)	1 = north, 2 = middle, 3 = local, 4 = south, 5 = east
First-semester credits (FSCRED)	1 = < 18 credits, 2 = [19, 22], 3 = > 23 credits
First-semester grade (FSGRADE)	1 = < 59.9, 2 = [60.0, 74.9], 3 = [75.0, 84.9], 4 = > 85.0
Second-semester credits (SSCRED)	1 = < 18 credits, 2 = [19, 22], 3 = > 23 credits
Second-semester grade (SSGRADE)	1 = < 59.9, 2 = [60.0, 74.9], 3 = [75.0, 84.9], 4 = > 85.0
Tuition fee waiver (FWAIVER)	0 = without tuition fee waiver, 1 = with tuition fee waiver
LOAN	0 = without loan, 1 = with loan
ABSENTEEISM	0 = < 10 classes, 2 = [11, 20], 3 = > 21
OUTCOME	0 = dropout, 1 = persist

Results

Descriptive statistics. Fourteen variables were tested by cross-tabulation and correlation tests and entered into the proposed models as a result of their statistical significance regarding OUTCOME. The six predictors obtained were: major of study (MAJOR), student age (AGE), place of residence (RESIDENCE), first-semester credits earned (FSCRED), second-semester grade (SSGRADE), and absence in class (ABSENTEEISM). The descriptive statistics of student retention are shown in Table 2.

Table 2. Descriptive Statistics of Student Retention

Variable		# of Dropouts	Dropout Rate (%)	# of Persisted	Persistence Rate (%)	Total
<i>MAJOR</i>	Engineering	51	13.1	339	86.9	390
	Business	28	12.8	191	87.2	219
	Social Science	29	14.5	171	85.5	200
	Security Science	8	5.5	138	94.5	146
<i>AGE</i>	< 18	8	28.6	20	71.4	28
	19–21	49	9.6	463	90.4	512
	> 21	59	14.2	356	85.8	415
<i>RESIDENCE</i>	North	28	20.6	108	79.4	136
	Middle	34	15.2	190	84.8	224
	Local	40	8.2	447	91.8	487
	South	11	12.4	78	87.6	89
	East	3	15.8	16	84.2	19
<i>FSCRED</i>	< 17	52	17.2	251	82.8	303
	18–21	55	11.9	408	88.1	463
	> 22	9	4.8	180	95.2	189
<i>SSGRADE</i>	< 59.9	61	53.0	54	47.0	115
	60.0–74.99	24	10.0	216	90.0	240
	75.0–84.99	23	5.8	374	94.2	397
	> 85.0	8	3.9	195	96.1	203
<i>ABSENTEEISM</i>	< 10	69	8.6	732	91.4	801
	11–20	18	17.0	88	83.0	106
	> 21	29	60.4	19	39.6	48
Total		116	12.1	839	87.9	955

The dropout rate was found to be 12.1 per cent, representing 116 student withdrawals before the second year of university. Students who had a major in Security Science, were aged between 19 and

22 years and were living at a local residence had a lower dropout rate than students from other majors, who were younger than 18 years or older than 21 years, and were not living locally.

Students with first-semester credits earned of greater than 22 credit points had a lower dropout rate. Second-semester grade and absenteeism had a significant influence on student dropout. The dropout rate decreased with increasing second-semester grade in the range of 3.9 to 53.0 per cent. Similarly, the dropout rate decreased with decreasing absenteeism in the range of 8.6 to 60.4 per cent.

Logistic regression. The regression model which included all the six predictors was used as the control model in this study. The other models were created using five predictors, excluding one variable in each model. The predictors included in each model are shown in Table 3.

Table 3 Model Predictors

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Control Model
<i>AGE</i>	v	v	v	v	v		v
<i>MAJOR</i> (<i>study major</i>)	v	v	v	v		v	v
<i>ABSENTEEISM</i> (<i>absence of class</i>)	v	v	v		v	v	v
<i>FSCRED</i> (<i>first- semester credit earned</i>)	v	v		v	v	v	v
<i>RESIDENCE</i>	v		v	v	v	v	v
<i>SSGRADE</i> (<i>second-semester grade</i>)		v	v	v	v	v	v

Based on the six predictors, seven regression models were created. The control model included all six predictors while Model 1 excluded second-semester grade, Model 2 excluded place of residence, Model 3 excluded first-semester credits earned, Model 4 excluded absenteeism, Model 5 excluded major of study, and Model 6 excluded age. Three statistics used to assess the models were: the exp (B) coefficient, log likelihood (log L), and the correct prediction rate. The statistics for the models are presented in Table 4.

Table 4 Logistic regression analysis for variables predicting dropout

Model	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Control Model
	exp(B)						
Major - Engineering	0.33	0.42	0.34	0.42	-	0.38	0.38
Major - Business	0.43	0.57	0.41	0.46	-	0.48	0.47
Major - Social Science	0.44	0.42	0.19	0.41	-	0.38	0.38
AGE	1.10	1.11	1.08	1.12	1.15	-	1.14
RESIDENCE							
- North	1.93	-	1.31	1.02	1.18	1.24	1.27
- Middle	2.56	-	2.01	1.88	1.91	2.04	2.05
- Local	5.56	-	4.06	3.72	3.93	4.20	4.29
- South	2.56	-	2.01	1.76	1.81	1.96	1.98
FSCRED	1.93	2.15	-	2.42	2.41	2.26	2.29
SSGRADE	-	3.33	3.33	4.02	3.47	3.42	3.43
ABSENTEEISM	0.29	0.62	0.59	-	0.66	0.63	0.63
Log Likelihood	457.96	412.29	410.36	403.26	402.46	398.71	398.37
Per cent correct prediction	62.7	62.8	65.6	66.9	69.7	71.1	71.1

The exp (B) statistic indicates the change in likelihood of dropping out of higher education associated with a one-unit change in the predictor variable. The higher the value of exp (B), the greater will be the effect of the predictor on the persistence rate. In Table 4, the highest value of exp (B) in the control model is reported for local place of residence which was 4.29. This indicates that local residents would make the greatest contribution to the persistence rate. With an exp (B) value of 3.43 for second-semester grade and 2.29 for first-semester credits earned, these two variables also have a great influence on the persistence rate in the control model. This indicates that academic performance, second-semester grade and first-semester credit earned affect student persistence. On the other hand,

local place of residence also has the highest value of exp (B) from Models 1 to 6, ranging from 3.93 to 5.56. This also confirms that local residence has a significant influence on student persistence. The second highest value of exp (B) is for second-semester grade from Models 2 to 6, ranging from 3.33 to 4.02. This is the second highest value of exp (B) from Models 2 to 6. This indicates that these two predictors—residence and second-semester grade—also have significant effects on dropout in Models 1 to 6.

The prediction rate identifies the best model for predicting persistence. The higher the prediction rate, the higher will be the accuracy of the model for predicting dropout behaviour. The accuracy of prediction of the models ranges from 62.7 per cent to 71.1 per cent. Model 1 has the lowest accuracy among the seven models studied. This shows that without second-semester grade the model cannot predict effectively. Model 6, which excludes age has the same accuracy of prediction rate, 71.7 per cent, with control model. It indicates that age does not have much influence on dropout as the other variables do. The control model has the best prediction accuracy, which is 71.1 per cent. This indicates that a better prediction model would require all six predictors.

Based on these findings, place of residence, second-semester grade and first-semester credits earned are reported to be three significant predictors of student persistence. On the other hand, the variable age is not as important as other variables for determining persistence.

Conclusion

Student retention modeling was performed on a dataset containing the records of 955 first-year students who were enrolled at a higher education institution in the 2013 academic year in Taiwan. The dropout rate was found to be 12.1 percent. Cross-tabulation and correlation tests performed that attributes, major field of study, age, type of residence, second-semester grade, first-semester credits, and absenteeism, were found to significantly determine students' dropout rates. Logistic regression results revealed that the control model which included all the six predictors had the highest level of prediction accuracy. The model without second-semester grade appears to be the worst model to predict dropout among the seven models. In addition, residence and first-semester credit were found to significantly influence dropout. It is recommended that better orientation programs and new programs for improving student academic performance be developed to enhance the retention rates.

References

- [1] R. Bennett, Determinants of undergraduate student dropout rates in a university business studies department, *Journal of Further and Higher Education*. 27(2003)123-141.
- [2] V. Tinto, Stages of Student Departure: Reflections on the Longitudinal Character of Student Leaving, *The Journal of Higher Education*. 59(1998)438-455.
- [3] V. Johnston, Why do first year students fail to progress to their second year? An academic staff perspective, presented at the British Educational Research Association Annual Conference (1997)
- [4] C. McInnis, From marginal to mainstream strategies: responding to student diversity in Australian universities. *European journal of education*, 38(2003)387-400.
- [5] K. E. Bigelow, Student perspectives in an all-female first-year Engineering Innovation course. *International Journal of Engineering Education*, 28(2012)286-292.
- [6] E. Ribeiro, G. Feixas, A.Maia, J. Senra, G. Dada, Changes in the Construct Systems of First-Year University Students: Impact on Psychological Symptoms and Problem-Solving Skills. *Journal of Constructivist Psychology*, 25 (2012) 170-180.