

Online Versus Face-to-Face: A Quantitative Study of Factors Influencing Students' Choice of Study Mode using Chi-Square Test and Binary Logistic Regression

Sze-Kiu Yeung

Singapore University of Social Sciences
Singapore

Wee Leong Lee

Singapore University of Social Sciences
Singapore

For online learning in the January 2019 semester, students at the Singapore University of Social Sciences were able to choose whether they want to study in virtual or face-to-face mode in two courses. Virtual refers to full online learning whereby students study, in a six-week term, without the need to meet face-to-face with the instructor while face-to-face refers to blended e-learning whereby students received either six or three face-to-face lessons with e-learning resources. In full online mode, students will meet the instructor virtually via video conferencing on a weekly basis. Data were obtained to find out which variables actually had an effect of students' choice of learning mode. 370 students were analysed and the variables including gender, marital status, race, nationality, course, qualification, school, programme, intake, age (now), age (joint) and cumulative grade point average (CGPA) were examined. Each variable was compared with the students' mode of study in order to identify if they are dependent (e.g. gender versus study mode, race versus study mode, etc.) based on a chi-square test. The significant variables were further investigated using a binary logistic regression model. It was found that qualification, intake and CGPA were found to be significant for students' choice of learning.

Keywords: Virtual, Face-to-face, Mode of study, Chi-square test, Binary logistic regression

Introduction

Two courses, Customer Relationship Management (BUS354) and Starting and Managing a Business (BUS357), offered by the School of Business provided two modes of study for students to undertake in the semester of January 2019. One was virtual whereby the students learned online with virtual face-to-face interactions with their instructors and peers while the other being physical face-to-face based on the blended e-learning approach of combining either six or three face-to-face lessons with e-learning contents over a term of six weeks. Students taking these courses were studying part-time taking classes in the evening and they had the option to choose their mode of study. All students taking these courses will take a common examination at the end, but the continuous assessment components will be different. Data from the Student Information Systems provided students' background information including demographic and academic details. 370 students were analysed and the variables that was extracted would include gender, marital status, race, nationality, course, qualification, school, programme, intake, age (now), age (joint) and cumulative grade point average (CGPA). The purpose of this paper was to find out from the data if there were significant variables that influenced students' choice of study. Insights drawn from this study will be helpful in planning for course offering in various modes. We believed the statistical analyses of the chi-square test and the binary logistic regression would be appropriate to obtain the findings for this study.

Literature Review

Online and face-to-face learning have been studied widely. Researchers had found that for online students they are usually older, have full or part-time work, requires commuting to the campus, have family obligations and have taken online courses before. Cleveland, Dutcher & Epps (2015) explained in their study that "online students tend to be older, part or full time workers, and returning to school after being in the working world" while the students in their survey who took the face to face "tended to be the more traditional college student: younger, often directly out of high school" (p. 128). On the other hand, face-to-face students are usually freshmen and they like to seek interactions with their instructors and classmates in the physical classroom. Dendir (2016) found that "the average online student was a sophomore, whereas the typical face-to-face student was a freshman a closer look at the data shows that 83% of the sample in the face-to-face section were freshmen, whereas about 77% in the online sample were sophomore and above a majority of the online students (58%) had prior experience with online courses" (p. 62). The key to online study is the flexibility and convenience to learn at the students' own pace and when they are most productive as pointed out by Jagers (2014) "that convenience and flexibility

are key factors that entice students to enroll in online coursework” (p. 27). In terms of student characteristics, it was found that “student age, percentage female, race and grade point average (GPA)” had no differences by the mode of delivery (Parcel, Radu & Gonzales, 2018, p.4). This study, based on two courses that allowed students to choose between either online or face-to-face mode, attempts to determine which independent variables affected students’ choice of learning.

Research Question

Based on the independent variables obtained for this study, would there be a significance between each independent variable compared to the mode of learning (i.e. the dependent variable) for the students who studied in BUS354 and BUS357? Would there also be interactions between these independent variables?

Chi-square Test and the Binary Logistic Regression Model

The use of the chi-square test and the binary logistic regression model as statistical tests came about from papers discussing the analysis of dependent variable in binary form. They included the studies of integrated pest management (IPM) adoption (Talukder, Sakib & Islam, 2017), drivers’ reactions in car crashes (Al-Taweel, Young & Sobhani, 2016) and stillbirths in Ethiopia (Berhie & Gebresilassie, 2016). These papers analysed the binary nature of the dependent variable (see Table 1) against a range of independent variables.

Table 1: Dependent Variables in Binary Format

Dependent Variable	Category
IPM Adoption	1: Yes, 0: No
Drivers’ Reactions	1: Drivers take Reactions, 0: Drivers do not take Reactions
Experienced Stillbirth	1: Yes, 0: No

Given that all the independent variables are in categorical format, the use of the chi-square test to determine the significance of the variables with the binary dependent variable made statistical sense. To further identify the levels of each independent variable such that there is significance associated with the dependent variable, these papers suggested the use of the binary logistic regression model. For example, are there significance associations between IPM adoption and different regions (Talukder et al., 2018), drivers’ reactions and crash type (Al-Taweel et al., 2016) and experiencing stillbirth and maternal age (Berhie & Gebresilassie, 2016). The key question for each study was to determine if there was IPM adoption, drivers take reactions or experiencing stillbirth among different levels of independent variables. From these studies, it was established that a consistent statistical approach using the chi-square test and the binary logistic regression model to determine if students’ choice of learning (virtual or face-to-face) was significant against a selection of independent variables would be valid. Details about the binary logistic regression model are explained in the journals from Peng, Lee & Ingersoll (2002) and Sperandei (2013).

Interpreting Results from the Binary Logistic Regression Model

In terms of interpreting the results of the binary logistic regression model, an understanding on the use of the odds ratio (OR) is important (Strand, Cadwallader & Firth, 2011). By definition, an OR ‘compares the odds of success (or failure) for a particular group to a base (reference) category for that variable’ (Strand et al., 2011, p. 18). For example, if we evaluate ethnicity and higher academic results according to Table 2, we note that White British students have been selected as the reference category. Indian students are 1.58 times more likely to achieve higher academic results than White British students. Conversely for Black Caribbean students the OR is 0.53, so Black Caribbean students are less likely to achieve higher academic results compared to White British students. In percentage terms they are 47% less likely to achieve higher academic results. What this means is that Indian students are more likely while Black Caribbean are less likely compared to White British students on achieving higher academic results. In SPSS, OR is represented by the ‘Exp(B)’ ratio.

Table 2: Ethnicity and Higher Academic Results

Ethnicity	OR (for Higher Academic Results)
0 White British	Reference Category
1 Mixed Heritage	0.87
2 Indian	1.58
3 Pakistani	0.64
4 Bangladeshi	0.80
5 Black Caribbean	0.53
6 Black African	0.81
7 Others	1.21

With an understanding on interpreting the results of the binary logistic regression model, it was possible to determine which levels of the independent variables were significant and were more likely or less likely compared to the reference category with respect to students' mode of study (i.e. virtual or face-to-face).

Methodology

The status of students' mode of study (1 for virtual, 0 for face-to-face) was considered as the main variable of interest (i.e. the dependent variable). If the student had chosen to study in full online learning, then he/she was considered as a virtual student otherwise face-to-face. At the same time, a list of other demographic characteristics (also considered variables) for students was also captured. These independent variables included gender, marital status, nationality, course, school, race, qualification, programme, intake, age (now), age (joint, i.e. when joining SUSS) and cumulative grade point average (CGPA). The question of interest was to determine which of these independent variables affected students' mode of study in terms of learning in virtual or face-to-face mode and, if possible, the interactions among them. Table 3 provides a listing of the variables and its category.

Table 3: Variables Selected in Relation to Students' Mode of Study

Variable	Category
Dependent	
Study Mode	1: Virtual, 0: Face-to-face
Independent	
Gender	1: Male, 2: Female
Marital Status	1: Single, 2: Married or Divorced
Nationality	1: Singaporean, 2: Others
Race	1: Chinese, 2: Malay, 3: Indian, 4: Others
Age (Now)	1: < 30, 2: 30 to 39, 3: >= 40
Age (Joint)	1: < 30, 2: 30 to 39, 3: >= 40
Course	1: BUS354, 2: BUS357
School	1: Business, 2: Others
Programme	1: BSBZ, 2: BSMA, 3: Others
Qualification	1: Diploma, 2: A-levels, 3: Others
Intake	1: 2012/01 to 2014/07, 2: 2015/01 to 2017/07, 3: 2018/01 to 2019/01
CGPA	1: <= 2.00, 2: 2.01 to 3.00, 3: > 3.00

Statistical Analysis

To assess students' mode of learning in either virtual or face-to-face mode, a statistical analysis involving univariate, bivariate and multivariate setup were conducted. In the bivariate setup, a chi-square test was used to assess the significance between the dependent and independent variables. In the multivariate setup, the binary logistic regression model was used to determine the likelihood of study mode (virtual or face-to-face) with the above independent variables that were found to be significant from the chi-square test. The software that was used for the analysis of data is SPSS (Version 22 for Windows).

Results and Discussions

Univariate Analysis

Among the 370 students, 42% were male and 58% female. The average age of students at intake was 26 and the age groups at intake were broken into three categories: less than 30 (82.2%), 30 to 39 (11.1%) and equal or higher than 40 (6.8%). The majority was single (81.1%) with married or divorced at 18.9%. In terms of race and nationality, the majority was Chinese (79.7%) with Malay (10.5%), Indian (6.2%) and Others (3.5%) as the other races and Singaporean (97.6%) was the majority nationality with the minority being other nationalities (2.4%). More students have studied BUS357 (57.8%) compared to BUS354 (42.2%) and they mostly studied in the School of Business (87.3%) compared to other schools (12.7%). The programmes that the students studied were evenly distributed between Bachelor of Business (BSBZ) (44.5%) and other programmes (42.8%) with the remaining students taking Bachelor of Marketing (BSBM) (12.7%). Qualification-wise, the majority of these students had a Diploma (73.2%) with A-levels (8.1%) and others (18.6%) as the remaining qualifications. For semester-intakes they were broken down into three periods: January 2012 to July 2014 (10%), January 2015 to July 2017 (75.4%) and January 2018 to January 2019 (14.6%). In term of the students' Cumulative Grade Point Average (CGPA) they were based on three ranges: less than or equal to 2.00 (10%), 2.01 to 3.00 (36.5%) and higher than 3.00 (53.5%). Finally, for the dependent variable of mode of study, the majority of students had chosen to study 'face-to-face' (87.6%) with the remaining students opted for 'virtual' (12.4%). Table 4 shows a summary of these statistics.

Table 4: Summary Statistics

Variable	Frequency	Percent
Gender		
Male	155	41.9
Female	215	58.1
Age (Joint)		
<30	302	82.2
30 to 39	41	11.1
>=40	25	6.8
Marital Status		
Single	300	81.1
Married or Divorced	70	18.9
Race		
Chinese	295	79.7
Malay	39	10.5
Indian	23	6.2
Others	13	3.5
Nationality		
Singaporean	361	97.6
Others	9	2.4
Course		
BUS354	156	42.2
BUS375	214	57.8
Qualification		
Diploma	271	73.2
A-Levels	30	8.1
Others	69	18.6
School		
Business	323	87.3
Others	47	12.7
Programme		
BSBZ	165	44.6
BSBM	47	12.7
Others	158	42.7
Intake		
2012/01 to 2014/07	37	10.0
2015/01 to 2017/07	279	75.4
2018/01 to 2019/01	54	14.6

CGPA		
<= 2.00	35	10.0
2.01 to 3.00	135	36.5
> 3.00	198	53.5
Study Mode		
Virtual	46	12.4
Face-to-Face	324	87.6

Bivariate Analysis

Based on the chi-square test of independence for categorical variables, the following hypotheses were evaluated:

- Null Hypothesis (H0): the two categorical variables were independent (i.e. there was no relationship between them);
- Alternative Hypothesis (H1): the two categorical variables were not independent (i.e. there was a relationship between them and that they were significantly related).

H0 was rejected if the Pearson chi-square had p-value < 0.05 or 0.10, meaning that statistically the two variables were significant.

As shown in Table 5, the independent variables that were reportedly to be highly significant ($p < 0.05$ or 0.01) or significant ($p < 0.10$) with the dependent variable (i.e. study mode) were qualification, programme, intake, age (joint) and CGPA. The chi-square test showed that these five independent variables were correlated with the dependent variable. On students' qualification, those with A-levels (26.7%) had opted to study in 'virtual' more than those with diploma (12.2%) and other qualifications (7.2%). The marketing programme (23.4%) was more popular with 'virtual' students compared to business (11.5%) and other programmes (10.1%). Intake-wise, the latest semester-intakes of January 2018 to January 2019 had a higher proportion (27.8%) of 'virtual' students compared to the first semester-intakes (January 2012 to July 2014) (21.6%) and the second semester-intakes (January 2015 to July 2017) (8.2%). For students' age (joint), those who were '40+' had the highest percentage of virtual learners (24.0%) compared to those who were aged 'less than 30' (12.5%) and between '30 to 39' (4.9%). As for the CGPA, the proportion of 'virtual' students were higher in the 'less than or equal to 2.00' group (29.7%) compared to '2.01 to 3.00' (16.3%) and 'greater than 3.00' (6.6%). On the other hand, the other seven independent variables (gender, marital status, race, nationality, course, school and age (now)) do not have significant effect on students' mode of study given that the Pearson chi-square test was $p > 0.10$.

Table 5: Assessing Association between Study Mode and Independent Variables with P values from Chi-square Test

Independent Variable	Study Mode (n = 370)		P Value
	Virtual (%)	Face-to-Face (%)	
Gender			0.581
Male	21 (13.5)	134 (86.5)	
Female	25 (11.6)	190 (88.4)	
Marital Status			0.905
Single	37 (12.3)	263 (87.7)	
Married or Divorced	9 (12.9)	61 (87.1)	
Race			0.891
Chinese	38 (12.9)	257 (87.1)	
Malay	5 (12.8)	34 (87.2)	
Indian	2 (8.7)	21 (91.3)	
Others	1 (7.7)	12 (92.3)	
Nationality			0.368
Singaporean	44 (12.2)	317 (87.8)	
Others	2 (22.2)	7 (77.8)	
Course			0.279
BUS354	16 (10.3)	140 (89.7)	
BUS375	30 (14.0)	184 (86.0)	
Qualification			0.026**
Diploma	33 (12.2)	238 (87.8)	
A-Levels	8 (26.7)	22 (73.3)	
Others	5 (7.2)	64 (92.8)	
School			0.941
Business	40 (12.4)	283 (87.6)	
Others	6 (12.8)	41 (87.2)	
Programme			0.047**
BSBZ	19 (11.5)	146 (88.5)	
BSBM	11 (23.4)	36 (76.6)	
Others	16 (10.1)	142 (89.9)	
Intake			0.000***
2012/01 to 2014/07	8 (21.6)	29 (78.4)	
2015/01 to 2017/07	23 (8.2)	256 (91.8)	
2018/01 to 2019/01	15 (27.8)	39 (72.2)	
Age (Now)			0.527
<30	35 (12.7)	240 (87.3)	
30 to 39	5 (8.6)	53 (91.4)	
>=40	6 (16.2)	31 (83.8)	
Age (Joint)			0.073*
<30	38 (12.5)	266 (87.5)	
30 to 39	2 (4.9)	39 (95.1)	
>=40	6 (24.0)	19 (76.0)	
CGPA			0.000***
<= 2.00	11 (29.7)	26 (70.3)	
2.01 to 3.00	22 (16.3)	113 (83.7)	
> 3.00	13 (6.6)	185 (93.4)	

***p value < 0.01, **p value < 0.05, *p value < 0.10

Binary Logistic Regression Analysis

We used the binary logistic regression model with the significant variables found in the bivariate analysis (see Table 5) to determine students' mode of study. The results are shown in Table 6. We noted that the independent variables of qualification ($p = 0.098 < 0.1$), intake ($p = 0.025 < 0.05$) and CGPA ($p = 0.007 < 0.01$) had significant effect on students' mode of study. For qualification, students with A-levels were almost twice (188%) more likely to take up virtual learning (OR = 2.875) compared to diploma holders (reference category). Students from semester-intakes of January 2015 to July 2017 were less likely (59%) to take virtual than those from the January 2012 to July 2014 (reference category) (OR = 0.407). In terms of students' CGPA, those with 'greater than 3.0' was less likely (77%) to take virtual than those from the 'less than or equal to 2.00' group (reference category) (OR = 0.228). However, programme and age (joint) were found not to have significant effect ($p > 0.1$) on students' mode of study even though it did have significant association from the bivariate analysis.

Table 6: Binary Logistic Regression Model for Students' Mode of Study (Virtual)

Dependent Variable	Independent Variable	Level	Coefficient	Odds Ratio (Exp(B))	p-value	
Mode of Study (Virtual)	Qualification	Diploma	Reference Category		0.098*	
		A-Levels	1.056	2.875	0.034**	
		Other	0.030	1.030	0.961	
	Programme	BSBZ	Reference Category			0.145
		BSBM	0.696	2.007	0.143	
		Other	-0.278	0.757	0.522	
	Intake	2012/01 to 2014/07	Reference Category			0.025**
		2015/01 to 2017/07	-0.899	0.407	0.072*	
		2018/01 to 2019/01	0.160	1.174	0.772	
		Age (Joint)	< 30	Reference Category		0.121
			30 to 39	-1.227	0.293	0.114
		>= 40	0.624	1.867	0.250	
	CGPA	<= 2.00	Reference Category			0.007***
		2.01 to 3.00	-0.501	0.606	0.293	
		> 3.00	-1.479	0.228	0.004***	

***p value < 0.01, **p value < 0.05, *p value < 0.10

In terms of the interactions between the independent variables, they were analysed according to, firstly, Age (Joint) by Marital Status, and Age (Joint) by Gender, since we thought female students who were at child bearing age or who were raising a family would choose virtual learning. Secondly, other independent variables were randomly selected to obtain the interaction results and they included Age (Joint) by CGPA, Age (Joint) by Qualification, Intake by Gender and CGPA by Qualification. Except for Intake by Gender, all interactions were not significant (see Table 7).

Table 7: Interaction between Independent Variables

Independent Variable	p-value
Qualification	0.971
Programme	0.063
Intake	0.023
Age (Joint)	1.000
CGPA	0.139
Age (Joint) * Marital Status	1.000
Age (Joint) * Gender	1.000
Age (Joint) * CGPA	0.993
Age (Joint) * Qualification	1.000
Intake * Gender	0.019
CGPA * Qualification	0.976

Conclusion

The chi-square test revealed that the independent variables of qualification, programme, intake, age (joint) and CGPA were found to be significant ($p < 0.10$) and that the bivariate analysis showed students have the highest percentage of virtual learning for A-levels (26.7%), marketing (23.4%), intake between January 2018 to January 2019 (27.8%), age (joint) of '40+' (24.0%) and CGPA 'less than or equal to 2.00' (29.7%). At the same time, the binary logistic regression model was used to analyse the adjusted effect of the levels from these independent variables (see Table 6). For qualification, A-levels holders were more likely (OR = 2.875) to take virtual learning compared to students with diploma (reference category) by 188%. On the other hand, students from the 'January 2015 to July 2017' intake were less likely (OR = 0.407) to take virtual learning compared to 'January 2012 to July 2014' (reference category) by 59% and those with CGPA of 'higher than 3.00' were less likely (OR = 0.228) against the reference category of 'less than or equal to 2.00' by 77%. For programme and age (joint), together with the other seven independent variables from the chi-square test (gender, marital status, race, nationality, course, school, age (now)), they had no significant effect on students' mode of study ($p > 0.5$). In conclusion, the levels that had an effect on students' choice of virtual learning were A-level holders, those admitted between January 2012 to July 2014, and students with CGPA of 'less than or equal to 2.00' while the interaction effects were minimum. Compared to the literatures this study confirmed that older students (based on intake data) had a preference for virtual learning. In terms of the paper's usefulness to learning and teaching, for older students in virtual learning a balanced use of technology (i.e. not overwhelming) would be suitable. The limitation of this study was that they were based on two courses. Data on current work experience would also be helpful since part-time students would work either part-time or full-time. Also, students' experience of other online courses would be helpful as well. In short a data-set with these additional variables would be more helpful but for this study the variables obtained are based on those shown in Table 3. Given that these two courses were offered as both 'virtual' and 'face-to-face' formats, this study, based on the data available, determined what variables influenced students' mode of learning according to the findings from the chi-square test and the binary logistic regression.

References

- Al-Taweel, H. M. H., Young, W., & Sobhani, A. (2016). A Binary Logistic Regression Model for the Driver Avoidance Manoeuvres in Tow Passenger Vehicle Crashes. *Australasian Transport Research Forum (ATRF)*, 38, 2016.
- Berhie, K. A., & Gebresilassie, H. G. (2016). Logistic Regression Analysis on the Determinants of Stillbirth in Ethiopia. *Maternal Health, Neonatology and Perinatology* (2016) 2:10. <https://doi.org/10.1186/s40748-016-0038-5>
- Cleaveland, M. C., Dutcher, C., & Epps, K. (2015). Comparing Business Law in Online and Face-to-Face Formats: A Difference in Student Learning Perception. *Academy of Educational Leadership Journal*, Vol 19 (1), 2015.
- Dendir, S. (2016). An online premium? Characteristics and performance of online versus face-to-face students in Principles of Microeconomics. *Journal of Education for Business* (2016), 91:2, 59-68, <https://doi.org/10.1080/08832323.2015.1110555>
- Jaggars, S. S. (2014). Choosing Between Online and Face-to-Face Courses: Community College Student Voices. *American Journal of Distance Education* (2014), 28:1, 27-38, <https://doi.org/10.1080/08923647.2014.867697>
- Parcel, T. L., Radu, M. B., & Gonzales, L. F. (2018). Who Selects an Online Class Over the Same Course Face-to-Face? And Who Learns More? Results from a Mixed-Methods, Quasi-Experimental Study of Teaching the Sociology of Work and Family. *The Journal of Public and Professional Sociology* (2018): Vol. 10: Issue 1, Article 6.

Peng, C. J., Lee, K. L., & Ingersoll, G. M. (2002). An Introduction to Logistic Regression Analysis and Reporting. *The Journal of Education Research*. Vol 96 (1), 2002. <https://doi.org/10.1080/00220670209598786>

Sperandei, S. (2014). Understanding Logistic Regression Analysis. *Biochemia Medica* 2014; 24(1):12-8. <https://doi.org/10.11613/BM.2014.003>

Strand, S., Cadwallader, C., & Firth, D. (2011). Using Statistical Regression Methods in Education Research. ESRC National Centre for Research Methods, University of Southampton, 2011. Retrieved from <http://www.restore.ac.uk/srme/www/fac/soc/wie/research-new/srme/index.html>

Talukder, A., Sakib, M. S., & Islam, M. A. (2017). Determination of Influencing Factors for Integrated Pest Management Adoption: A Logistic Regression Analysis. *Agrotechnology* 6: 163. <http://dx.doi.org/10.4172/2168-9881.1000163>

Please cite as: Yeung, S.K. & Lee, W.L. (2019 Online Versus Face-to-Face: A Quantitative Study of Factors Influencing Students' Choice of Study Mode using Chi-Square Test and Binary Logistic Regression. In Y. W. Chew, K. M. Chan, and A. Alphonso (Eds.), *Personalised Learning. Diverse Goals. One Heart. ASCILITE 2019 Singapore* (pp. 340-348).