*International Education Journal*, 2007, 8(1), 194-204.
ISSN 1443-1475 © 2007 *Shannon Research Press*.
http://iej.com.au


# Measuring student course evaluations: The use of a loglinear model

**Ding Hooi Ting**
School of Business, Monash University, Malaysia *ting.ding.hooi@buseco.monash.edu.my*

**Mireya Sosa Abella**
Universiti Utara, Malaysia *misoabel3@yahoo.com*

*In this paper, the researchers attempt to incorporate the marketing theory (specifically the service quality model) into the education system. The service quality measurements have been employed to investigate its applicability in the education environment. Most of previous studies employ the regression-based analysis to test the effectiveness of course evaluations and teaching. In econometric term, the use of regression-based analysis contains some bias. This is because it does not capture the complex and interaction terms of the variables under study. In this article, the authors try to compare the use of a loglinear model versus the widely used normal-regression analyses to evaluate the students' course evaluations of Spanish classes at Universiti Utara Malaysia. SERVQUAL model has been employed in order to examine the usefulness of the model under study. The findings from the study shows that the loglinear model provides a better analytical procedure for students' course evaluation as it is able to explain more variation in the model under study.*

SERVQUAL, regression analysis, Spanish classes, course evaluation, loglinear model

## INTRODUCTION

In the academic world, it is important to evaluate the effectiveness of courses conducted from students' perspectives. Hence the use of a good analytical tool is important in order to capture the variance explained by the independent variables in the model under study. Previous studies have used extensively the regression-based models to explain the behaviour of the subject matter under study. These studies looked only at the linear relationship between the dependent variable and the independent variable (s). However, those studies were not able to capture the fundamental variations involved in the study. This is because the use of regression based analysis is not able to capture all the interaction terms among the variables, and it is crucial to use analytical tools that are able to capture interaction effects among the variables, since the interaction terms may be able to explain the variations involved in the model in greater detail. Hence this study is employed to evaluate the applicability of the loglinear model in an academic setting, specifically in the Spanish courses in Universiti Utara Malaysia (UUM). The Spanish courses were used as the focal-point of the study because it is a language course; hence, there was need to have extra interactions between instructor and students and also a better means of course delivery.

The objective of the study is to compare the use of a loglinear model versus the widely used normal regression analysis to investigate the students' course evaluations on Spanish classes at Universiti Utara Malaysia (UUM). The independent variables used in this study are basically from the service quality (SERVQUAL) measurements while the dependent variable used involves the overall satisfaction level of students taking the Spanish classes.

## LITERATURE REVIEW

Goods are objects. A service is a deed and a performance. When goods are being purchased, something tangible is acquired. Berry et al. (1988) argued that because of the intangibility of services, customers evaluated service quality based on the tangible environment. LeBlanc et al. (1988) and Zeithaml et al., (1985) asserted that the problem of evaluating service quality was more difficult and complicated compared to product quality because of its intangibility. Since services were intangible in nature and the inseparability of production and consumption, it would be difficult for customers to perform prior evaluation of a firm's services; hence the trustworthiness, believability and credibility of the service provider were crucial determinants of patronage (Malhotra et. al. 1994). How a service was performed would affect the customers' perception of the quality of the services rendered.

Services delivered are difficult to evaluate because they are delivered to people by people and cannot be standardized because of their heterogeneous nature (Gupta et al., 1988; LeBlance et al., 1988; Rushton et al., 1989; Zeithaml et al., 1985). The same contact person is not able to give the same level of services to another customer even though he is serving the next customer immediately. Furthermore, different providers have different methods of giving services and the service level also differs from time to time.

Many academicians give different definitions of quality. Customers' expectations are the true standard for judging service quality and not the policy of the bank or the management of the bank (Berry et al., 1988). Parasuraman et al. (1991) pointed out that customers expected service companies to do what they were supposed to do (fundamentals), not fanciness; performance and not empty promises. Therefore, defining customer needs in the service industries is more complex compared to the manufacturing industries because the customers are involved in the production process.

Berry et al (1988) argued that customers assessed service quality by comparing what they wanted or expected to what they were getting. Customers would build expectations of services based on earlier experiences, communication, image, word-of-mouth and the customers' need (Holmlund et al., 1996). Quality was how the offer of the bank gained uniqueness and value in the eyes of the customers and it was both the act of making the offer different and its evaluation by customers (Christopher et. al., 1994). The customers would be the judges in determining the success of the services rendered.

Westbrook (1980) defined satisfaction as the subjectiveness of individual evaluations. While Bearden et al. (1983) defined satisfaction as the positive final outcome when using the scarce resources. From the above definitions, satisfaction was related to the subjective emotional evaluation (Andreassen, 2000).

According to Yuksel et al. (1998), determining customer satisfaction was fundamental towards delivering effective services. Commitment to customer satisfaction must be an ongoing process because regardless of how well service was provided, customers would always push for higher levels (Murray, 1991).

## METHOD OF ANALYSIS

This study employs the SERVQUAL measurement introduced by Parasuraman et al., (1988, 1994) (see Appendix) to measure the quality of service and the satisfaction level of the students in the Spanish classes at UUM. SERVQUAL is the abbreviation for service quality. SERVQUAL is represented by five dimensions (Parasuraman et al. 1988); tangibles, reliability, responsiveness, empathy and assurance and is measured using the differences between the scales for perceptions and expectations. Perceptions are defined as customers' beliefs concerning the services received

(Parasuraman et al., 1988). Whereas expectations are viewed as predictions made by customers about what is likely to happen during certain encounters (Parasuraman et al., 1988), namely, they do not represent predictions about what service providers "would" offer but what they **should** offer (Teas, 1993). This model has been used extensively in the West, but most of the researchers have used the normal regression-based analysis. In this study, another approach has been employed to study the relationship between the students' course evaluations and the service quality using the loglinear model which is able to capture the interaction terms.

Students from the Spanish courses Level I to Level III were chosen as the respondents. This study included all respondents taking the Spanish courses, as the number of students taking the courses was limited.

## VARIABLES OF STUDY

There is evidence that suggests that evaluation of satisfaction should involve a curvilinear or higher order form as well as an interaction effect (Taylor and Baker, 1997; Ting, 2004). This was confirmed by Oliva, Oliver and MacMillan (1992) who stated that the satisfaction function should not be in the linear form. A study by Basadur and Head (2001) suggested that a non-linear relationship might exist for satisfaction where they looked at the effectiveness of teamwork in relation to organizational success. Hence, these findings more or less falsified the methods used in past research studies, that were the normal regression based analyses.

From the academic point of view, students' assessments of professors and courses and the interpretations of the statistical analyses used in the past researches had long interested many educators (Mehdizadeh, 1990). A major problem with previous works was the analysis technique, namely, the widely used regression analysis. More recent studies in education have taken into consideration the deficiencies of normal regression or the ordinary least squares models and employed the logit models. In comparison to the loglinear models, logit models are easier to formulate, but they do not incorporate the most general interaction terms among the independent variables. In a sense, logit models disregard the complex structural relationships that may exist among the independent variables. Hence, Mehdizadeh (1990) has introduced the usefulness of the loglinear models to analyse the students' course evaluations.

There have been many studies on the possible factors influencing students' evaluation of teaching effectiveness (Dilts, 1980; Seiver, 1983) and students' course evaluations (Kelly, 1972). According to Seiver (1983), expected grade on the last day of class could be a good indicator of a students' course evaluation. Fizel and Johnson (1986) stated that the current GPA would serve as an indicator of a student's attitude and academic performance and it was expected to have a positive effect on a student's attitude. On top of that, the SERVQUAL model was also included as discussed in the literature review.

Each of the variables was included because their expected effects upon students course evaluation had already been established (Mehdizadeh, 1991). This would help to validate the use of loglinear models.

## QUESTIONNAIRES

A set of questionnaires developed by Parasuraman et al. (1988, 1994) was used to carry out the research. The questionnaire employed consisted of four parts as follows:

Part A: 5 questions on the background of respondents
Part B: 22 questions on expectation
Part C: 22 questions on performance
Part D: 2 open-ended questions

## RESPONDENTS BACKGROUND

Table 1 below shows the respondents' background in this study. The total number of respondents who took part in this study was 71 students.

**Table 1. Background of respondents**

| | | Frequency | Percentage |
|---|---|---|---|
| **Sex** | Male | 12 | 16.9 |
| | Female | 59 | 83.1 |
| **Race** | Malay | 25 | 35.2 |
| | Chinese | 29 | 40.9 |
| | Indians | 13 | 18.3 |
| | Others | 4 | 5.6 |
| **Semester** | 1 | 1 | 1.4 |
| | 2 | 2 | 2.8 |
| | 3 | 10 | 14.1 |
| | 4 | 1 | 1.4 |
| | 5 | 46 | 64.8 |
| | 6 | 3 | 4.2 |
| | 7 | 8 | 11.3 |
| **Class Obtained*** | First Class | 2 | 2.8 |
| | Second Class Upper | 41 | 57.8 |
| | Second Class Lower | 28 | 39.4 |
| **Entry** | Matriculation | 8 | 11.3 |
| | STPM | 62 | 87.3 |
| | Diploma | 1 | 1.4 |
| **Total** | | 71 | 100.00 |

* Class obtained is their previous semester's CGPA where 3.67 – 4.00 falls into first class, 3.00 – 3.66 falls into second class upper and 2.99 and below falls into second class lower.

Most of the respondents are female students. This scenario corresponds to UUM situation where most of the students are females. They comprise 83 per cent of the total respondents, while the male respondents are only 17 per cent. Most of the respondents are Chinese (41%) followed by Malays (35%), Indians (18%) and others (6%).

Most of the students are in their fifth semester that is 65 per cent of the total respondents. This is followed by respondents from Semester 3 (14%). The lowest number of respondents who took part in this study are respondents from Semester 1 (1%).

The respondents were asked to indicate the class that they obtained during their previous semester. Though they had yet to graduate, the class they obtained was important as it was used as one of the variables in this study. The majority of the respondents obtained second class upper (58%), while a large number managed to obtain second class lower (39%). Only two respondents obtained a first class degree.

The entry qualification of the respondents was from STPM (87%) and Matriculation (11%). There was only one respondent with a diploma.

## EMPIRICAL TESTING USING REGRESSION ANALYSIS

The first step in this analysis is to test the usefulness of normal regression analysis on the variables under study. In general terms, from the data collected the overall course evaluation is regressed on the independent variables (class, tangible, reliability, responsiveness, assurance and empathy). Table 2 reports the results obtained.

From Table 2, only two out of six parameters (including the constant term) are significant at $p < 0.01$ level while the rest are non-significant. The $R^2$ obtained is only 0.399 which shows that

only 40 per cent of the variance of the overall course evaluation can be explained by the independent variables (tangible, reliability, responsiveness, assurance and empathy), while 60 per cent is not explained by the model. The $\overline{R}^2$ value obtained is only 0.396 after taken into consideration the number of parameters in the study. Hence, the normal regression analysis results show that the model is not very useful in explaining the overall course evaluation of students.

**Table 2. Empirical results from regression analysis**

| | b | Std. Error | $R^2$ | $\overline{R}^2$ |
|---|---|---|---|---|
| (Constant) | 2.58** | 0.95 | 0.399 | 0.396 |
| Class | -0.25 | 0.22 | | |
| Tangible | -0.01 | 0.15 | | |
| Reliability | 0.02 | 0.24 | | |
| Responsiveness | -0.19 | 0.26 | | |
| Assurance | 0.76** | 0.21 | | |
| Empathy | 0.02 | 0.13 | | |

** significant at $p < 0.01$ level

The data were then transformed into categorical variables so that the loglinear regression analysis could be performed.

## TRANSFORMING THE VARIABLES INTO CATEGORICAL VARIABLES

In order to perform the loglinear model, all the variables were transformed into categorical variables as is seen in Table 3 which is denoted by value 1 and value 2 respectively. Since the Likert scale ranges from 1 to 7, the responses were condensed to average or below (scales that ranged from 1 to 4) and above average (scales that ranged from 5 to 7).

**Table 3. Collapsing of the categories into condensed classes**

| | | Value | Frequency | Percent |
|---|---|---|---|---|
| Y | **Overall** | | | |
| | Average or Below | 1.0 | 16 | 22.54 |
| | Above Average | 2.0 | 55 | 77.46 |
| $X_1$ | **Class** | | | |
| | Second Class Lower | 1.0 | 29 | 40.85 |
| | First of Second Class Upper | 2.0 | 42 | 59.15 |
| $X_2$ | **Tangible** | | | |
| | Average or Below | 1.0 | 13 | 18.31 |
| | Above Average | 2.0 | 58 | 81.69 |
| $X_3$ | **Reliability** | | | |
| | Average or Below | 1.0 | 7 | 9.86 |
| | Above Average | 2.0 | 64 | 90.14 |
| $X_4$ | **Responsiveness** | | | |
| | Average or Below | 1.0 | 5 | 7.04 |
| | Above Average | 2.0 | 66 | 92.96 |
| $X_5$ | **Assurance** | | | |
| | Average or Below | 1.0 | 9 | 12.68 |
| | Above Average | 2.0 | 62 | 87.32 |
| $X_6$ | **Empathy** | | | |
| | Average or Below | 1.0 | 9 | 12.68 |
| | Above Average | 2.0 | 62 | 87.32 |
| | **Total** | 71 | 71 | 100.00 |

From the table above, a large majority of the respondents agreed that the Spanish courses conducted were good. This was because most of them said that the lowest percentage value was 55 per cent and the highest percentage value was 66 per cent and that the courses conducted were above average.

## LOGLINEAR MODEL

Loglinear analysis of categorical data is not a new technique in data analyses Mehdizadeh (1990). It is not commonly used compared to the normal regression analysis. In a loglinear model, the number of hierarchical models that can be fitted into the model increases as the number of independent variables increases. Hence this makes the identification of a well-fitting unsaturated model more difficult (an unsaturated model contains fewer than all possible interactions of the dependent variables).

The concepts of loglinear modelling can be introduced in a more systematic way. The simplest model is a two-dimensional contingency table of Y to $X_1$. In a loglinear model of y to $x_1$, the number of cases in each cell can be expressed as the function of Y, $X_1$ and the interaction terms of $Y*X_1$. In order to obtain the loglinear model, the natural logs of the estimated cell frequencies are used. In general the expression of a loglinear model is as follow:

$$\text{Ln } f_{ij} = \alpha_i^y + \alpha_j^x 1 + \alpha_{i*j}^{y*x} 1 + \varepsilon$$

Loglinear models allow researchers to investigate the interrelation among many variables with complex structures by using the conditional independence. Thus, all effects of independent variable upon a response variable are revealed rather than just the marginal effect of each independent variable.

The t-tests are used to test the significant value of the lambdas. If the test of interdependence of Y and $X_1$ is of interest, then the independence model $\text{Ln } f_{ij} = u + \lambda_i^y + \lambda_j^y$ can be examined. The model does not incorporate the interaction term, where the $f_{ij}$ is no longer the observed frequency in the (i, j)th cell but the expected frequency based on the model. An iterative algorithm is used to obtain the estimates of the lambda parameters.

In order to obtain a linear model, the natural logs of the estimated expected frequencies rather than the actual counts are used. The observed frequencies and the estimated expected frequencies for fitting the independence model to the cross-classification of students' course evaluations and their expected grades and the natural logs of the estimated expected cell frequencies are shown in Table 4.

The two dimensional marginal table (Table 4) with a two-dimensional cross-classification have been examined.

From Table 4, the values obtained were then calculated in order to find the estimated λ parameters for students' overall course evaluations and the class obtained.

Table 5 contains the estimates of the λ parameters for the main effects and their interactions for all cells of Table 5. The interpretations of the obtained λ values are simple and it is similar to the interpretations of the main effects and interaction effects in a usual analysis.

In order to examine whether the loglinear model (the interaction effects) plays a significant role, the interaction effects and the main effects are compared as in Table 5. Since in general, the interaction effects are greater than the main effects, hence further analysis is conducted.

Further analysis was carried out in order to estimate the Pearson chi-square and the likelihood ratio chi-square statistics for each model (Table 6). The first model is the simplest model with no interactions. The second model includes the interaction between y and $x_0$ because a student's overall service quality evaluations and his or her expected grade seem to be interrelated- (Y) ($X_0$) ($X_1$) ($X_2$) ($X_3$) ($X_4$) ($X_5$) ($YX_0$). All the models seem to be significant at $p < 0.001$.

**Table 4. Observed frequencies, estimated expected frequencies and the natural logs of the estimated cell frequencies for independence model of students' overall course evaluations and class**

| Students' Overall Course Evaluations | Class | | Total (Averages) |
|---|---|---|---|
| | Second Lower | First and Second class upper | |
| ≤ average | 10[a] (6.4)[b]<br>2.3026 (1.8563) | 6 (3.6)<br>1.7918 (1.2809) | 16 (1.5686)[e] |
| > average | 18[c] (21.6)[d]<br>2.8904 (3.0727) | 36 (32.40)<br>3.5835 (3.4782) | 54 (3.2754) |
| Total (Averages) | 28<br>(2.4645)[e] | 42<br>(2.3796) | 70 (2.4220) |

[a] The observed frequencies
[b] The estimated frequencies under the independence model
[c] The natural logs of the observed frequencies
[d] The natural logs of the estimated frequencies
[e] The averages of the natural logs of the estimated frequencies

Note: To estimate the expected frequency of each cell under the model of independence of row and column variables and under the model of binomial proportions (homogeneity of proportions) in the two-dimensional tables, one may use the following formula:

$$\hat{m}_{ij} = \frac{x_{it}\ x_{tj}}{N} \quad where \quad i = 1, 2 \quad and \quad j = 1, 2$$

where $x_{it}$ and $x_{tj}$ are the sums of row and column frequencies, N is the sample size in the table, and $\hat{m}$ is the estimated expected frequency of the cell under the independence model.

**Table 5. Estimated of λ parameters for students' overall course evaluations and their class**

$\lambda_1^y = 1.5686 - 2.4220 = -0.8534$
$\lambda_2^y = 3.2754 - 2.4220 = 0.8534$
$\lambda_1^{x_1} = 2.4645 - 2.4220 = 0.0425$
$\lambda_1^{x_2} = 2.3796 - 2.4220 = -0.0424$
$\lambda_{1*1}^{y*x_1} = 2.3026 - (2.4220 - 0.8534 + 0.0425) = 0.6915$
$\lambda_{2*1}^{y*x_1} = 2.8904 - (2.4220 + 0.8534 + 0.0425) = -0.4275$
$\lambda_{1*2}^{y*x_1} = 1.7918 - (2.4220 - 0.8534 - 0.0424) = 0.2656$
$\lambda_{2*2}^{y*x_1} = 3.5835 - (2.4220 + 0.8534 - 0.0424) = 0.3505$

In order to test the difference in the $G^2$ when new interaction terms are included, the final table (Table 6) is produced.

Based on the final results obtained in Table 7, the best model is:

$(Y)(YX_0)(YX_1)(YX_2)(YX_3)(YX_4)(YX_5)(X_0X_1)$, where the $G^2$ is 0.727.

Though the $G^2$ keeps on increasing after the inclusion of other interaction terms, but the significant change showed a non-significant value at $p < 0.05$ after that. Furthermore, the $G^2$ change is too small (only differs by 0.015 from its subsequent model) and not worthwhile to include further terms in the model as the additional contribution can be omitted. From the comparison between the normal regression based analyses as in Table 2, the variation explained for normal regression analysis was only 40 per cent compared to 73 per cent using the loglinear model. Hence, this study supports the claim that the loglilnear model is superior to the normal based regression analysis.

**Table 6. Values of the Chi Square Goodness-of-Fit Statistics, $\chi^2$ and $G^2$ for various Loglinear Models Concerning Students Overall Course Evaluations**

| Model | df | $\chi^2$ | p | $G^2$ | p |
|---|---|---|---|---|---|
| (Y) ($X_0$) ($X_1$) ($X_2$) ($X_3$) ($X_4$) ($X_5$) | 57 | 594.848 | 0.000 | 83.605 | 0.000 |
| ($YX_0$) | 124 | 762.593 | 0.000 | 302.242 | 0.000 |
| ($YX_1$) | 124 | 271.587 | 0.000 | 494.380 | 0.000 |
| ($YX_2$) | 124 | 476.881 | 0.000 | 257.526 | 0.000 |
| ($YX_3$) | 124 | 463.748 | 0.000 | 249.040 | 0.000 |
| ($YX_4$) | 124 | 495.087 | 0.000 | 260.545 | 0.000 |
| ($YX_5$) | 124 | 459.840 | 0.000 | 259.994 | 0.000 |
| ($X_5X_0$) | 124 | 587.515 | 0.000 | 277.798 | 0.000 |
| ($X_5X_1$) | 124 | 372.905 | 0.000 | 243.883 | 0.000 |
| ($X_5X_2$) | 124 | 362.848 | 0.000 | 227.202 | 0.000 |
| ($X_5X_3$) | 124 | 356.738 | 0.000 | 212.430 | 0.000 |
| ($X_5X_4$) | 124 | 369.948 | 0.000 | 233.020 | 0.000 |
| ($X_4X_0$) | 124 | 624.000 | 0.000 | 282.542 | 0.000 |
| ($X_4X_1$) | 124 | 400.912 | 0.000 | 243.858 | 0.000 |
| ($X_4X_2$) | 124 | 372.134 | 0.000 | 217.353 | 0.000 |
| ($X_4X_3$) | 124 | 370.842 | 0.000 | 218.087 | 0.000 |
| ($X_3X_0$) | 124 | 585.094 | 0.000 | 265.361 | 0.000 |
| ($X_3X_1$) | 124 | 382.814 | 0.000 | 234.979 | 0.000 |
| ($X_3X_2$) | 124 | 360.549 | 0.000 | 211.931 | 0.000 |
| ($X_2X_0$) | 124 | 597.564 | 0.000 | 272.449 | 0.000 |
| ($X_2X_1$) | 124 | 377.868 | 0.000 | 234.426 | 0.000 |
| ($X_1X_0$) | 124 | 639.279 | 0.000 | 291.007 | 0.000 |
| ($YX_5X_4$) | 120 | 285.083 | 0.000 | 208.960 | 0.000 |
| ($YX_5X_3$) | 120 | 264.902 | 0.000 | 193.452 | 0.000 |
| ($YX_5X_2$) | 120 | 281.883 | 0.000 | 207.600 | 0.000 |
| ($YX_5X_1$) | 120 | 279.200 | 0.000 | 219.478 | 0.000 |
| ($YX_5X_0$) | 120 | 438.254 | 0.000 | 253.887 | 0.000 |
| ($YX_4X_3$) | 120 | 283.555 | 0.000 | 198.069 | 0.000 |
| ($YX_4X_2$) | 120 | 273.055 | 0.000 | 194.593 | 0.000 |
| ($YX_4X_1$) | 120 | 298.768 | 0.000 | 215.473 | 0.000 |
| ($YX_4X_0$) | 120 | 461.182 | 0.000 | 196.212 | 0.000 |
| ($YX_3X_2$) | 120 | 286.600 | 0.000 | 196.212 | 0.000 |
| ($YX_3X_1$) | 120 | 291.390 | 0.000 | 211.964 | 0.000 |
| ($YX_3X_0$) | 120 | 448.058 | 0.000 | 245.330 | 0.000 |
| ($YX_2X_1$) | 120 | 291.473 | 0.000 | 213.791 | 0.000 |
| ($YX_2X0$) | 120 | 442.825 | 0.000 | 248.883 | 0.000 |
| ($YX_1X_0$) | 120 | 471.495 | 0.000 | 264.327 | 0.000 |

Note: $G^2$ is the chi-square likelihood ratio.
y = overall Students' Course Evaluations;
$X_0$ = class; $X_1$ = tangible; $X_2$ = reliability; $X_3$ = responsive; $X_4$ = assurance; $X_5$ = empathy

## LIMITATIONS

The first limitation appears when different individuals foresee qualities from different aspects. This is because service qualities are very subjective. What is perceived as good quality by one individual may not be good for another individual.

The second limitation is other students who have taken the subject earlier may influence the perceptions of service quality and satisfaction of the present students. This is because some students have the tendency to ask their seniors about the subject or course and the information given by them may have a significant impact on their perceptions.

**Table7. Partitioning of the Likelihood Ratio Statistics for Partial-Association Test**

| Model | $G^2$ | $\Delta G^2$ | Sig. |
|---|---|---|---|
| $(YX_0)$ | 0.352 | 0.352 | 0.000 |
| $(YX_0)(YX_1)$ | 0.563 | 0.210 | 0.000 |
| $(YX_0)(YX_1)(YX_2)$ | 0.602 | 0.039 | 0.013 |
| $(YX_0)(YX_1)(YX_2)(YX_3)$ | 0.652 | 0.050 | 0.003 |
| $(YX_0)(YX_1)(YX_2)(YX_3)(YX_4)$ | 0.653 | 0.002 | 0.591 |
| $(YX_0)(YX_1)(YX_2)(YX_3)(YX_4)(YX_5)$ | 0.655 | 0.002 | 0.570 |
| $(YX_0)(YX_1)(YX_2)(YX_3)(YX_4)(YX_5)(X_0X_1)$ | 0.727 | 0.072 | 0.000 |
| $(YX_0)(YX_1)(YX_2)(YX_3)(YX_4)(YX_5)(X_0X_1)(X_0X_2)$ | 0.742 | 0.015 | 0.061 |
| $(YX_0)(YX_1)(YX_2)(YX_3)(YX_4)(YX_5)(X0X_1)(X_0X_2)(X_0X_3)$ | 0.754 | 0.012 | 0.084 |
| $(X_0X_3)(YX_1)(YX_5)(YX_2)(YX_0)(YX_3)(YX_4)(X_0X_1)(X_0X_2)(X_0X_4)$ | 0.757 | 0.003 | 0.462 |
| $(X_0X_3)(YX_1)(YX_5)(YX_2)(YX_0)(YX_3)(YX_4)(X_0X_1)(X_0X_2)(X_0X_4)(X_0X_5)$ | 0.762 | 0.002 | 0.462 |
| $(X_0X_3)(YX_1)(YX_5)(YX_2)(YX_0)(YX_3)(YX_4)(X_0X_1)(X_0X_2)(X_0X_4)(X_0X_5)(X_1X_2)$ | 0.762 | 0.000 | 0.787 |
| $(X_0X_3)(YX_1)(YX_5)(YX_2)(YX_0)(YX_3)(YX_4)(X_0X_1)(X_0X_2)(X_0X_4)(X_0X_5)(X_1X_2)(X_1X_3)$ | 0.762 | 0.000 | 0.734 |
| $(X_0X_3)(YX_1)(YX_5)(YX_2)(YX_0)(YX_3)(YX_4)(X_0X_1)(X_0X_2)(X_0X_4)(X_0X_5)(X_1X_2)(X_1X_3)(X_1X_4)$ 0.764 | 0.002 | 0.500 | |
| $(X_0X_3)(YX_1)(YX_5)(YX_2)(YX_0)(YX_3)(YX_4)(X_0X_1)(X_0X_2)(X_0X_4)(X_0X_{5)(}X_1X_2)(X_1X_3)(X_1X_4)(X_1X_5)$ | 0.765 | 0.001 | 0.731 |
| $(X_1X_5)(X_0X_3)(YX_4)(YX_5)(YX_1)(X_0X_2)(X_0X_5)(YX_0)(X_1X_2)(X_0X_1)(X_1X_4)(X_1X_3)(X_0X_4)(YX3)(YX_2)(X_2X_3)$ | 0.775 | 0.012 | 0.090 |
| $(X_1X_5)(X_0X_3)(YX_4)(YX_5)(YX_1)(X_0X_2)(X_0X_5)(YX_0)(X_1X_2)(X_0X_1)(X_1X_4)(X_1X_3)(X_0X_4)(YX_3)(YX_2)(X_2X_3)(X_2X_5)$ 0.786 | 0.009 | 0.144 | |
| $(X_1X_5)(X_0X_3)(YX_4)(YX_5)(YX_1)(X_0X_2)(X_0X_5)(YX_0)(X_1X_2)(X_0X_1)(X_1X_4)(X_1X_3)(X_0X_4)(YX_3)(YX_2)(X_2X_3)(X_2X_{5)(}X_3X_4)$ | 0.788 | 0.002 | 0.456 |

Y = Overall Students' Course Evaluation
$X_0$ = Class Obtained
$X_1$ = Tangible
$X_2$ = Reliability
$X_3$ = Responsive
$X_4$ = Assurance
$X_5$ = Empathy

## DIRECTION FOR FUTURE RESEARCH

In this research analysis, there is the view that the loglinear model is able to provide greater predictive power for the data under study when compared with the regression-based analysis. Hence, further studies need to be done in order to support this view and the findings in other settings.

Furthermore, future research studies need also to look for other factors that may contribute to the variations in the course evaluations such as the duration of contact, the quality of the text-book used, assignments related to the course and the applicability of the contents of the course under study.

## CONCLUSIONS

This article shows the usefulness of loglinear models to analyse categorical variables when compared to the normal regression analysis. The approach involves fitting a general hierarchical model to a data set drawn from student overall course evaluations. The findings from this analysis show that students' class, tangibles, reliability, responsiveness, assurance and empathy tend to interact with each other, contributing to a higher $G^2$ value. Hence the usage of the loglinear model that captures the interaction terms can be said to explain students' course evaluation better than the ordinary regression analysis. Through this research, it is hoped that new insights can be gathered in order to serve better the students in the coming semesters and the areas that need to be focused on, while conducting the Spanish classes in UUM.

## REFERENCES

Andreassen, T.W. (2000). Antecedents To Satisfaction With Service Recovery. *European Journal of Marketing*, 34 (1/2), 156-175.

Basadur, M. and Head, M. (2001). Team Performance and Satisfaction: A Link to Cognitive Style within a Process Framework. *Journal of Creative Behavior*, 34 (4), 227-248.

Bearden, W. D. and Teel, J. E. (1983). Selected Determinants of Customer satisfaction, and Complaint Reports. *Journal of Marketing Research*, 20 (November), 21 – 28.

Berry, L.L., Parasuraman, A. and Zeithaml, V.A. (1988). The Service-Quality Puzzle, *Business Horizons*, (September-October), 35-43.

Christopher, M., Payne, A. and Ballantyne D. (1994). *Relationship Marketing-Bringing Quality, Customer Service, and Marketing Together.* Butterworth Heinemann.

Dilts, D. A. (1980). A statistical interpretation of student feedback. *Journal of Economic Education*, 11 (Spring), 10-15.

Fizel, J.L. and Johnson, J.D. (1986). The effect of macro/micro course sequencing on learning and attitudes in principles of economics. *Journal of Economic Education*, 17 (Spring), 87-98.

Gupta, Yash P. and Torkzadeh, G. (1988). Re-designing bank service systems for effective marketing. *Long Range Planning*, 21 (6), 38-43.

Holmlund, M. and Kock, S. (1996). Relationship Marketing: The Importance of Customer-Perceived Service Quality in Retail Banking. *The Service Industries Journal*, 16 (3), 287-304.

Kelly, A.B. (1972). Uses And Abuses Of Course Evaluation As Measures Of Educational Output. *Journal of Economic Education*, 4 (Fall), 13–18.

LeBlanc, G. and Nguyen, N. (1988). Customers' perception of service quality in financial institutions. *International Journal of Bank Marketing*, 6 (4), 7 - 18.

Malhotra, Naresh K. and Ulgado, Francis M. (1994). A Comparative Evaluation Of The Dimensions Of Service Quality Between Developed And Developing Countries. *International Services Marketing*, 11 (2), 5-15.

Mehdizadeh, M. (1990). Loglinear models and student course evaluations. *Research in Economic Education*, 21 (1), 7-21.

Murray, K.B. (1991). A Test of Services Marketing Theory: Customer Information Acquisition Activities. *Journal of Marketing*, 55 (January), 10-25.

Oliva, T. A., Oliver, R. L. and MacMillan, I.C. (1992). A Catastrophe Model for Developing Service Satisfaction Strategies. *Journal of Marketing,* 56 (3), 83-95.

Parasuraman, A, Zeithaml, V. and Berry, L.L (1994). Alternative scales for measuring service quality: a comparative assessment based on psychometric and diagnostic criteria. *Journal of Retailing*, 70 (3), 201-229.

Parasuraman, A., Berry, L.L. and Zeithaml, V.A. (1991). Understanding Customer Expectations Of Service. *Sloan Management Review*. 39 (Spring), 39-48.

Parasuraman, A., Zeithaml, V. A. and Berry, L. L. (1988). SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64 (1), 12-40

Rushton, Angela M. and Carson, David J. (1989). The marketing of services: managing the intangibles. *European Journal of Marketing*, 23 (8), 23-44.

Seiver, D. A. (1983). Evaluations and grades: a simultaneous framework. *Journal of Economic Education*, 14 (Summer), 32-38.

Taylor, Steven A. (1997). Assessing Regression-Based Importance Weights for Quality Perceptions and Satisfaction Judgements in the Presence of Higher Order and/or Interaction Effects. *Journal of Retailing*, 73 (1), 135-159.

Teas, K. (1993). Expectations, performance evaluation and consumers' perception of quality. *Journal of Marketing*, 57, 18-34.

Ting, D. H. (2004). "Service Quality and Satisfaction Perceptions: Curvilinear and Interaction Effect," International Journal of Bank Marketing, 22 (6), 407 – 420.

Westbrook, R.A. (1980). Intrapersonal Affective Influences Upon Consumer Satisfaction with Products. *Journal of Consumer Research*. 7 (June), 49-54.

## APPENDIX: SERVQUAL[1]

**Tangibles**
Having modern-equipment equipment/up-to-date equipment
Visually appealing facilities
Appearance of lecturer
Visually appealing materials associated with the course

**Reliability**
Keeping promise to do something by a certain time
Sincere interests in solving students' problems
Teach correctly the first time
Provide their services at the time they promised to do so
Insist on error- free records

**Responsiveness**
Tell students when services will be performed
Give prompt services to students
Always willing to help students
Respond to students' requests

**Assurance**
Instill confidence in students
Feel safe with the lecturers
Consistently courteous with students
Knowledge to answer students' questions

**Empathy**
Give students individual attention
Consultation hours convenient to all their students
Give students personal attention
Have the students' best interest at heart
Understand the specific needs of their students

**Overall Course Evaluation[2]**
My feelings towards the overall course can be best as

[1] Responses were obtained on seven-point scales ranging from "strongly disagree" to "strongly agree"
[2] Responses were obtained on seven-point scales ranging from "very dissatisfied" to "very satisfied"