Romblon State University Research Journal
ISSN: 2619-7529 (Online) | ISSN: 2350-8183 (Print)
Volume 3 (2): 1-7, 2021

Forecasting Enrollment: A Substitute to Career Guidance Campaign

Joefel T. Libo-on[1,2], and Al M. Gadon[1,3]

ABSTRACT

The use of forecasting models is essential for advanced decision-making and planning. The conduct of the career guidance campaign of guidance counselors and teachers has been a challenging task in a time of community and border restriction. This study was conducted to determine factors that can predict and forecast the number of Junior High School enrollees based on the eleven-year historical records. The data were gathered from the school registrar, guidance counselor of Tanagan National High School, and Philippine Statistics Authority (PSA). The unemployment rate, sex, parents' educational attainment, honors, the distance of the school from home, parent's income, employment of parents, and government beneficiaries are found to be significant predictors of enrollment. There were two forecasting models found and by confirmatory factor analysis (CFA) by structural equation modeling (SEM), the most appropriate model is $\hat{Y} = -1.714 + .335x_4$; where Y refers to the number of enrollment and x_4 to parents' educational attainment. Experimental results yield an average of 1.26% difference in the forecasted and actual numbers of enrollment. The forecasted enrollment using the most appropriate model underwent validation, and the model forecast was accepted.

Keywords: *AMOS, CFA, Enrollment, forecasting, SEM*

INTRODUCTION

The state of national emergency drives the Philippines to undergo community restriction to counter the spread of the virus, COVID-19. In such an event, the President of the Republic declares the whole Luzon archipelago under enhanced community quarantine on March 16, 2020 (Merez, 2020). The declaration of community restriction impacts every sector and affects the economy. This is the time where the education sector is on the culminating part of the school year for most of the basic education and mostly in the middle of the term for the tertiary and graduate education level.

The surge of COVID-19 cases also impacts the programs and activities of schools and universities. Among these is the conduct of career guidance. Expecting an enrollment based on the conduct of career guidance to the school within the catchment area is done when the school year is about to end. Thus, the challenge now is to determine a method to deliver career guidance that will consider the implemented safety guidelines of the national government to avoid face-to-face contact with students and the slow internet connection.

Enrollment is an event that happens every year in schools. Secondary schools are conducting an early enrollment campaign for incoming Junior High School students. This is to estimate the number of incoming students which may serve as a basis for the planning and preparation of schools. If the enrollment can be estimated as early as possible it will also help teachers to prepare an estimated number of copies of modules to be reproduced during the state of national emergency. Unfortunately, after enrollment, data are being kept aside purposelessly. However, the said data can be used to predict the number of enrollees for the near future using forecasting, which is a process of predicting something to happen using the variables of a certain event (Rakhimov & Kankarej, 2015).

Moreover, if the school can predict the number of enrollees by forecasting, it can lead to better plans and preparations (Lavilles & Arcilla, 2012). The school could prepare if there is a need to add classrooms, teachers, and other resources. Furthermore, if the school is well-prepared because of the foreseen possible number of enrollees, it can offer quality education which may produce equally competitive graduates.

For several years, Tanagan National High School has been interviewing enrollees and collecting other requirements of the incoming junior high school

✉ onemanbrave@gmail.com

[1]Graduate Education and Professional Studies, Romblon State University
[2]Admission Office, Romblon State University
[3]Tanagan National High School

Received 9 September 2021; Accepted 03 March 2022; Available online 7 April 2022

students. After this, data are just kept until the admission process resulting in a large amount of data in databases and student files in the registries. However, these are not just simply data; this could serve as a source of information that could help for the improvement and preparation of the school, especially during this time of community restrictions.

To make these data worthy, the researcher decided to utilize the gathered valuable information. By forecasting these data and processing those variables, the researcher could get some information like predicting the number of enrollees for the next school year and determining what factors contribute to students' decision to enroll in this school. Daniel and colleagues (2014), define forecasting as the process of making a statement on unobserved events and sometimes kept for estimates of values at a certain future period. Meanwhile, Lazar and Lazar (2015) stated that forecasting students' enrolled every year represents an important activity for an academic institution. This has been the basis of identifying the income and the basis of the foundation for operational plans and strategies. Not only does forecasting the data could help the school by predicting the number of enrollees but also be the basis of all plans and preparations.

Tanagan National High School experiences a lack of classrooms due to the sudden increase in the population of students every year. Way back year 2012 and below, the school only needed eight classrooms but at present, the school needs 13. However, the school has only 11 classrooms that is why the computer laboratory room and Technology and Livelihood Education (TLE) building were used as classrooms for the meantime just to accommodate the students. To add, the performance of the students, as well as the quality of education, is highly at stake because of the absence of conducive and enough classrooms.

This study seeks to provide a forecasting model of the possible number of enrollees for the succeeding years so that the school will have an idea about how many classrooms are still needed to be constructed for the next school years. Also, it aims to determine the factors that can predict the school enrollment rate. Moreover, if the school can already predict the possible number of students, the school management may have better plans and preparations especially in the reproduction of the number of modules. This involves proper planning of the resources which could not be possible if there will be no well-founded estimation of the students who are admitted annually (Lazar & Lazar, 2015).

METHODOLOGY

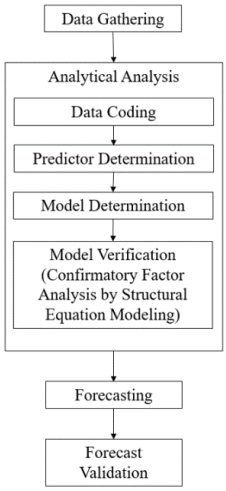

Figure 1. Paradigm of the study.

This study utilizes analytical method research. It is characterized by a phenomenon that answers the questions of how many. Description plays a critical role in educational research. The purpose of this analysis is to identify and explain patterns and differences in populations, establish new measures of key phenomena or describe samples in studies aimed at detecting causal effects (Loeb, et al., 2017).

This study was conducted at Tanagan National High School in Tanagan, San Andres, Romblon. The historical data of enrollment of Tanagan National High School from 2009 to 2019 were used in this study. These data were gathered from the school registrar, guidance counselor, and PSA with permission from the school principal. The paradigm for this study is presented in Figure 1. The data were gathered from the school registrar, guidance counselor, and PSA and were analyzed analytically. In an analytical analysis, the gathered data were analyzed, (Table 1) and the predictors and covariance were determined using correlation. Those independent variables that are significantly correlated are considered as predictors. Those that are not predictors will be dropped in the next step. In model determination, the determined predictors will be subjected to multiple regression to generate the available forecasting model given the determined predictors. After determining the forecasting models, it will be subjected to confirmatory factor analysis by structural equation modeling using analysis of moment structures to verify its model fitting.

Romblon State University Research Journal
ISSN: 2619-7529 (Online) | ISSN: 2350-8183 (Print)
Volume 3 (2): 1-7, 2021

Table 1. Correlation between variables.

		Number of Enrollees	Sex	Parents Educational Attainment	Honors	The distance of School from Home	Parents Income	Employment	Government Beneficiaries	Unemployment Rate	Inflation Rate	Attrition Rate	Catchment Area
Number of Enrollees	*r*	1											
	p												
Sex	*r*	**0.963	1										
	p	0.000											
Parents Educational Attainment	*r*	**0.994	**0.970	1									
	p	0.000	0.000										
Honors	*r*	**0.975	**0.953	**0.973	1								
	p	0.000	0.000	0.000									
The distance of School from Home	*r*	**0.919	**0.884	**0.888	**0.895	1							
	p	0.000	0.000	0.000	0.000								
Parents Income	*r*	**0.969	**0.950	**0.972	**0.959	**0.831	1						
	p	0.000	0.000	0.000	0.000	0.002							
Employment	*r*	**0.907	**0.834	**0.904	**0.872	**0.817	**0.874	1					
	p	0.000	0.001	0.000	0.000	0.002	0.000						
Government Beneficiaries	*r*	**0.945	**0.936	**0.948	**0.912	**0.851	**0.926	**0.768	1				
	p	0.000	0.000	0.000	0.000	0.001	0.000	0.006					
Unemployment Rate	*r*	*-0.620*	-0.519	-0.600	*-0.623	*-0.626	-0.584	-0.519	*-0.722	1			
	p	0.042	0.102	0.051	0.041	0.040	0.059	0.102	0.012				
Inflation Rate	*r*	-0.183	-0.118	-0.168	-0.046	-0.277	-0.167	-0.286	-0.180	0.233	1		
	p	0.590	0.729	0.622	0.894	0.410	0.623	0.395	0.597	0.491			
Attrition Rate	*r*	-0.290	-0.263	-0.253	-0.252	-0.445	-0.312	-0.068	-0.376	0.448	0.523	1	
	p	0.387	0.434	0.454	0.454	0.170	0.350	0.843	0.255	0.167	0.099		
Catchment Area	*r*	0.584	0.518	0.525	0.552	0.546	*0.613	0.408	*0.612	-0.530	0.070	-0.436	1
	p	0.059	0.103	0.097	0.079	0.082	0.045	0.212	0.046	0.094	0.837	0.180	

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

Table 2. ANOVA Table of Regression Model

	Model	Sum of Squares	df	Mean Square	F	Sig.	Interpretation	Decision
1	Regression	2759.184	1	2759.184	744.347	0.000[b]	Significant	Accept the Generated Model
	Residual	33.362	9	3.707				
	Total	2792.545	10					
2	Regression	2776.378	2	1388.189	686.884	0.000[c]	Significant	Accept the Generated Model
	Residual	16.168	8	2.021				
	Total	2792.545	10					

[a]*Dependent Variable: Number of Enrollees*
[b]*Predictors: (Constant), Parents Educational Attainment*
[c]*Predictors: (Constant), Parents Educational Attainment, Distance of School from Home*

The forecasting model with the best fit will be considered as the best forecasting model of the number of enrollments for Junior High School. Also, the best forecasting model will be utilized in predicting the estimated number of enrollees for the succeeding year. Lastly, the forecasted data generated by the best model will be validated. The statistical forecasting technique was utilized in forecasting junior high school enrollment. The technique made use of the data gathered for past periods to forecast future values. *Stepwise Regression Analysis (SRA)* was then utilized to pick out the potential explanatory factors that significantly contribute to the explanation of the criteria measure ($\hat{Y}$) – the number of enrollments. The process for SRA

included the computation of the coefficient of determination (R^2). *Adjusted R – Square (RA^2)* was computed to know if the independent variables included in the equation significantly contributed to the variation of the variable. Root Mean Square Residual (RMR) is the square root of the average squared amount by which the sample variances and covariances differ from their estimates obtained under the assumption that the model is correct. The smaller the RMR is, the better. An RMR of zero indicates a perfect fit. Comparative Fit Index (CFI) refers to the degrees of freedom and the non-centrality parameter estimate for the baseline model. The CFI is truncated to fall in the range from 0 to 1. CFI values close to 1 indicate a very good fit. The Root Mean Square Error of Approximation (RMSEA) model fit was used to give an overall idea if the structured model fitted.

Table 3. Regression Model

Dependent Variable	Predictors	Regression Equation	Adjusted R^2
Junior High School Enrollment	Parent's Educational Attainment	$\hat{Y} = -1.714 + 0.335x_4$	0.987
	Parent's Educational Attainment and Distance of School from Home	$\hat{Y} = 1.013 + 0.284x_4 + 0.064x_6$	0.996

RESULTS AND DISCUSSION

Predictors of Enrollment

The factors with a significant relationship or with predictive relationship to the number of enrollments is shown in Table 2. Furthermore, the significant covariance of each significant variance was considered during the Confirmatory Factor Analysis (CFA).

Model Identification

The stepwise regression was utilized to pick out from the independent variables the possible predictors or factors that could significantly affect the dependent variable. To determine the predictors of enrollment as a coded data, regression analysis was done with the aid of SPSS.

Model 1 is an acceptable model, F (1,9) = 744.347, p=0.000 (Table 2). By regression analysis, the parent's educational attainment predicted enrollment with the coefficient of determination of .994 (Table 2), calculating the proportion of variance of the enrollment over its mean (Table 4), which was explained by the parent's highest educational attainment, R^2 = .987. It implied that the parent's highest educational attainment provides 98.7% explanatory power to the number of enrollees. In addition, the majority of students' parents were only high school achievers therefore most of them have low income so, to send their children to school they will choose the nearest school. The mathematical model in predicting the number of enrollments was determined by:

$$\text{Model 1: } \hat{Y} = -1.714 + .335x_4$$

Model 2 is also another acceptable model, F(2,8)= 686.884, p=0.000 (Table 3). By regression analysis, the parent's highest educational attainment and distance of the school from home were identified as predictors of enrollment, R^2 = .996 or its adjusted coefficient of determination (Table 4) indicates that the number of students enrolled can be explained by the parent's highest educational attainment and distance of the school from home by 99.6% of its variation.

Table 4. CFA Statistics of $\hat{Y} = -1.714 + .335x_4$

Statistics	
AIC= 6.000	RMSEA= 0.630
CFI= 1.000	RMR= 0.000; PCLOSE=0.028

One of the factors was the distance of the school from home, which implies that the majority of the parents will send their children to the nearest school for low expenses. Although some came from isolated areas far from school, still they have no choice because the nearest school to them was Tanagan National High School. This is also an additional input to the result of the study of Li and Qiu (2018) that establishes parenting and the desire of parents to compete for better education quality. The mathematical model in predicting the number of enrollment was determined by:

$$\text{Model 2: } \hat{Y} = 1.013 + .284x_4 + .064x_6$$

Model Verification

Application of Analysis of Moment Structures (AMOS), the Structural Equation Modeling (SEM) for Confirmatory Factor Analysis (CFA) was advanced.

$$\textit{CFA for Model 1: } \hat{Y} = -1.714 + .335x_4$$

The explanatory powers of model structure of $\hat{Y} = -1.714 + .335x_4$ was affirmed by CFA as 0.987

(Figure 2); in the proportion of its variation, the number of enrollees over its mean was explained by parents' highest educational attainment.

Figure 2. Output-structured model for regression.

Statistics of the first structure model in Table 4 transpires using the independent variable described in model 1 shows estimates appear to be equivalent in both AIC (Akaike Information Criterion) = 6.000 The CFI (Comparative Fit Index) of 1 appears that the model is acceptable (Schrieber et. al, 2006). The root means square residual (RMR) of 0.000 indicates a highly acceptable model (Bian, 2011). Lastly, the RMSEA (Root Mean Square Error of Approximation) of 0.630 with *p*=0.028 indicates that the model is on its best fit.

CFA for Model 2: $\hat{Y} = 1.013 + .284x_4 + .064x_6$

Table 6. CFA Statistics of the Generated Model

Category	Model 1	Model 2
AIC (Akaike Information Criterion)	6.000	12.000
CFI (Comparative Fit Index)	1.000	1.000
RMR (Root Mean Square Residual)	0.000	0.000
RMSEA (Root *Mean* Square Error of Approximation) [PCLOSE]	0.630[.028]	0.407[.117]

The CFA affirms that the explanatory powers of model 2 was 0.996 (Figure 3); it implies that the enrollees was explained by parents' highest educational attainment and distance of the school from home. Model structure 2 (Table 6) estimates appear to be equivalent in both AIC = 12. The CFI of 1 appears that the model is acceptable (Schrieber et. al, 2006). The *RMR* = 0.000 indicates a highly acceptable model (Bian, 2011). Lastly, the *RMSEA*=0.407, *p*=0.117 indicates that the model is on its best fit.

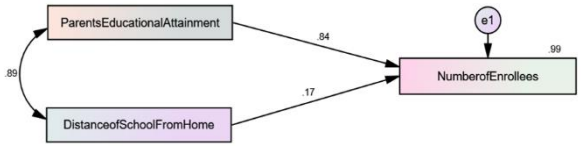

Figure 3. Output-structured model for regression.

Table 5. CFA Statistics of the $\hat{Y} = 1.013 + .284x_4 + .064x_6$

Statistics	
AIC= 12.00	RMSEA= 0.318
CFI= 1.000	RMR= 0.000; PCLOSE=0.117

Confirmatory Factor Analysis Comparative Summary Report

This section presents the best model in predicting the enrollment of Tanagan National High School. Since the models were all none nested models, the AIC appears that Model 1 appears the smallest estimates weights of 6.000 (Table 7) and considered as best fit among the other models (Schreiber et.al, 2006), *RMSEA*=0.630 (Figure 4).

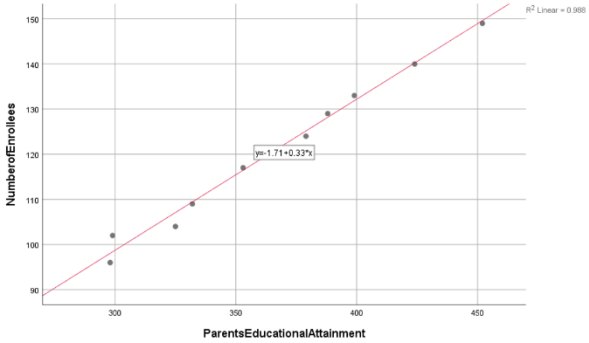

Figure 4. Graph of $\hat{Y} = -1.714 + .335x_4$

Forecasting Enrollment

Using Model 1 as the best fit model (Figure 5), the actual enrollment of the Junior High School was compared to the forecasted enrollment, and it shows an average of 98.74% accuracy or an average of 1.26% difference to the actual enrollment data.

Table 7. Acceptability of the Forecasted Data

Category	N	Mean	t	df	Sig. (2-tailed)	Interpretation	Decision
Actual	11	120.36	0.019	21	0.985	Not Significant	Accept the Forecasted Data

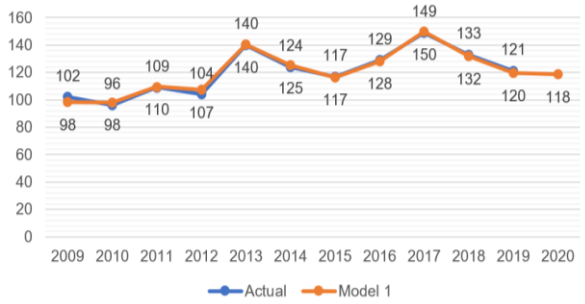

Figure 5. Comparison of the actual and forecasted enrollment.

Forecast Validation

Comparing the actual and the forecasted data generated by experimentation of model 1 shows that the forecasted data are somewhat the same, $t_{(21)} = 0.019$, p=0.985; hence, the forecasted data of the number of enrollments is acceptable.

CONCLUSION

The investigators concluded, based on the result, that the predictors of junior high school enrollment were unemployment rate, sex, parents' educational attainment, honors, the distance of the school from home, parent's income, employment of parents, and government beneficiaries. This added to the findings of previous studies where it was found that color and those with economic advantage are the predictors of student enrollment. Two equation models have been developed which shows that parent's highest educational attainment and distance of the school from home are significant predictors of enrollment. This indicates that the number of students to enroll can be predicted using forecasting models.

Parent's highest educational attainment was the predictor of the best model. This only shows that those students with parents of students having educational attainment affect the enrollment of Junior High School. The forecasted enrollment for the year 2020 was 118. The forecasted number of enrollments differs from the actual number of enrollments by 1.26%. This provides the school heads to opt that instead of conducting face-to-face career guidance in different schools, forecasting may be considered also avoiding a health-related risk during the pandemic. Generally, it is safe to conclude that majority of the parents of the upcoming enrollees of Tanagan National High School were educated.

AUTHORS' CONTRIBUTION

J.T.L.: Conceptualization, methodology, validation, formal analysis, writing original draft, supervision, revision. A.M.G.: Conceptualization, formal analysis, investigation, writing original draft, and revision.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Bian, H. (2011). *Structural Equation Modeling with AMOS II. https://www.coursehero.com/file/41626351/SEM-with-AMOS-IIpdf/*

Daniel, A.E., Norman, O.K., Chrisman, J.N., Ray, C.A., Smith, S.D. & Miller, M.W. (2014). *New WSR-88D operational techniques: responding to recent weather events.* https://ams.confex.com/ams/94Annual/webprogram/Paper241216.html

Lavilles, R. Q. & Arcilla, M. B. (2012). Enrollment forecasting for school management system. *International Journal of Modeling and Optimization 2* (5), 563-566.

Lazar, C., & Lazar, M. (2015). Forecasting Methods of the enrolled students' number economic insights. *Trends and Challenges, 4* (67), 41-51.

Li, Z., & Qiu, Z. (2018). How does family background affect children's educational achievement? Evidence from Contemporary China. *Journal of Chinese Sociology. 5* (13), 1-21.

Loeb, S., Morris, P., Dynarski, S., Reardon, S., McFarland, D., & Reber, S. (2017). Descriptive analysis in education: A guide for researchers. NCEE 2017-4023. *National Center for Education Evaluation and Regional Assistance.*

Merez, A. (2020). *Luzon under enhanced community quarantine as COVID-19 cases rise.* https://news.abs-cbn.com/news/03/16/20/duterte-places-luzon-

under-enhanced-community-quarantine-to-fight-covid-19

Rakhimov, I. & Kankarej, M. (2015). Forecasting the number of students in general education in university colleges using mathematical modeling. *Journal of Mathematical Sciences, Advances and Applications, 32*, 57–71.

Schreiber, J., Nora, A., Stage, F., Barlow E., & King, J. (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *Journal of Educational Research, 99*, 323-338.