PROFESSIONAL FILE

ARTICLE 129

Association for Institutional Research

A COMPARISON STUDY OF RETURN RATIO–BASED ACADEMIC ENROLLMENT FORECASTING MODELS

Xinxing Anna Zan, Sang Won Yoon, Mohammad Khasawneh, Krishnaswami Srihari

ACKNOWLEDGEMENTS

This study was supported by the Watson Institute for Systems Excellence at the State University of New York at Binghamton. The authors wish to thank Sandra Starke, Tania Das, and Jill Heneghan from the Institutional Research and Assessment Office for providing us with all the data for this research. We also thank anonymous reviewers who have given valuable comments to improve this paper.

About the Authors

Xinxing Anna Zan is a master of science student, Sang Won Yoon is assistant professor, Mohammad Khasawneh is associate professor, and Krishnaswami Srihari is distinguished professor. All are with the Department of Systems Science and Industrial Engineering, State University of New York at Binghamton, Binghamton, NY 13902.

The corresponding author, Sang Won Yoon, can be reached at the following: Phone: +1-607-777-5935; Fax: +1-607-777-4094, Email: yoons@binghamton.edu.

Abstract

In an effort to develop a low-cost and user-friendly forecasting model to minimize forecasting error, we have applied average and exponentially weighted return ratios to project undergraduate student enrollment. We tested the proposed forecasting models with different sets of historical enrollment data, such as university-, school-, and division-level enrollment. The numerical results indicate that the proposed models perform better when the school-level and university-level analyses are applied, as compared to the division-level analysis. We also observed that the forecasting error is lower when the most recent enrollment data sets are considered than when we consider all past enrollment data. In addition, when forecasting for spring semesters, the 1-year average return ratio method, using the school-level analysis, yields the lowest forecasting error of 0.40%. When forecasting for fall semesters, the average return ratio method, using the university-level analysis, yields the lowest forecasting error of 0.81%.

INTRODUCTION

Undergraduate student enrollment patterns require an accurate forecasting model to assist in college and university strategic planning efforts. Forecasting is the projection, estimation, or prediction of future event occurrences that are uncertain in nature (Tersine, 1994). Accurate forecasting can help people plan wisely for an organization’s future. Over the years, many forecasting models and techniques have been applied in business organizations, government agencies, educational systems, and public services (Armstrong, 2001). However, it is necessary to use suitable and accurate forecasting techniques. As a result, identifying the best enrollment forecasting model for a college or university is critical to effective decision making; such a forecasting model relates to the following services from a strategic planning perspective (Desjardins et al., 2006; Glover, 1986; Hossler & Bean, 1990; Norton, 1998):

- Improve the accuracy of student enrollment and income or budget forecasts

- Given the limitations of existing resources, offer high-quality academic programs as well as great campus experiences to meet students' needs
- Project campus housing assignments, necessary building or classroom planning, staffing allocation, and course scheduling
- Plan for total budgeting and allocate or balance income and expense at realistic levels for both academic values and student enrollment demands
- Seek ways to describe, analyze, predict, and improve student retention percentages
- Improve student and staff satisfaction
- Link and coordinate activities of recruitment, admissions, financial aid, and career planning

Therefore, the objective of this research is to develop a low-cost forecasting model to minimize forecasting error and to provide a less computationally intensive method that can be used to forecast undergraduate student enrollment. The research questions include the following:

1. How can we accurately forecast total undergraduate student enrollment in a college or university at a specific semester to reduce forecasting error?
2. How can we minimize the forecasting error when projecting undergraduate student enrollment (i.e., division-, school-, or university-level analyses)?
3. When projecting undergraduate student enrollment, should we apply all available past enrollment data or the most recent enrollment data for better forecasting accuracy?

One limitation of this paper is that our research does not consider the projection of freshman student enrollment, since it is typically given. Also, we make several forecasting assumptions to project student enrollment and improve forecasting accuracy.

The remainder of the paper is organized as follows. First, we present the background associated with this research. Then, we address the detailed research methodologies. Next, we explain and discuss numerical results of forecasting models. Finally, we summarize conclusions and opportunities for future work.

RESEARCH BACKGROUND

Many approaches and methods have been studied and proposed to forecast student enrollment, with each forecasting model generating different forecasting errors (Guo, 2002). Among other purposes, probability forecasting methods are used to calculate probabilities of an uncertain event happening in the future. A linear probability model was proposed where transition probabilities are used to calculate student enrollment (Marshall & Oliver, 1970). In their model, Marshall and Oliver considered students' total work to be done, based on probabilities of students attending, vacationing or interning, and dropping out. Marshall and Oliver applied this to forecast student enrollment at the University of California, Berkeley. Similar to the linear probability model, logit or probit models are used occasionally to forecast student enrollment when the outcome or dependent variables are known. However, logit or probit models are usually used to analyze more-complex educational behaviors (Porter, 1999).

Another category of probability forecasting methods is known as the "ratio forecasting method." For example, student retention rate has been utilized to forecast campus student enrollment for the University of Wisconsin-Madison (Beck, 2009). His results showed that the students' retention rate generally decreases as they approach their graduation year. Their results also showed that the decline of the students' retention rate decreased dramatically from 4–5 school years to 5–6 school years, largely due to graduation. We draw several important features from the ratio analysis (Beck, 2009):

- Models should be analyzed separately, especially when the retention ratios or percentages are significantly different.
- Retention patterns of students should be examined carefully based on historical data points and previous student enrollment data.
- Models should be tested against recent or past actual enrollment data.

The ratio method has also been applied to forecast student enrollment at the University of Washington (Schmid & Shanley, 1952) using a three-step procedure. First, they derived a series of estimates for the entire population for which all or at least a major component of the student enrollment is drawn. Therefore, they took the data for this category directly from the population forecast from the State of Washington. Second, it was important to learn the enrollment trends for all institutions of higher education in the State of Washington as a whole. This was necessary to determine the relationship between student enrollment and the age group between 18 to 21 years in the entire population during the past 30 years and during the next 10 years. Third, they determined the trend in the ratio of student enrollment to the total population of the age group between 18 and 21 years for the forecasting period. Based on detailed analysis of the historical data, they calculated the ratio between student enrollment and total population age group, and utilize that ratio for a future student enrollment projection. As a result, the proposed ratio method showed several advantages:

- Institutional researchers need expend less time and labor in performing student enrollment forecasts.
- Institutional researchers are able to use historical data to forecast.
- Institutional researchers do not need to define parameters or variables.

Furthermore, the exponential weighted moving average method is one type of moving average method that does not require a large amount of historical data points or records (Dobbs, 2001). The exponential weighted moving average methods vary, depending on single or multiple exponential smoothing approaches. The idea of the exponential weighted moving average method is that it smooths out variations in a time-series model by applying more weights on the more-recent data than on previous data (Tersine, 1994). Exponential weight forecasting methods have been used widely in operations research and economics (Muth, 1960). Many researchers used this forecasting method to predict short-term sales in inventory control (Brown, 1959; Magee, 1958). Nowadays, many other fields have started using the exponential weighted average forecasting method to predict different forecasts, because they have seen the success of this method in forecasting sales.

For instance, the double exponential smoothing method was studied as a pattern-based method to apply and adapt to a number of circumstances (Gardner, 1981). In this research, Gardner compared different forecasting methods, including correlation analysis, intention survey, and professional judgment methods to predict student enrollment where double exponential smoothing has the most reasonable and most consistent results. According to the literature (Dobbs, 2001; Holt, 2004; Snyder, 1988), the exponential weighted moving average has the following forecasting advantages:

- It includes all previous data to represent the entire history of data.
- It is easy to compute and provides better forecasting results for short-term projections.
- It does not require a large amount of historical data to implement.
- It provides flexibility in forecasting with seasonal behaviors and trends.
- In this research, therefore, the probability or ratio forecasting method, combined with the exponential weighted moving average method, is used in predicting university student enrollment.

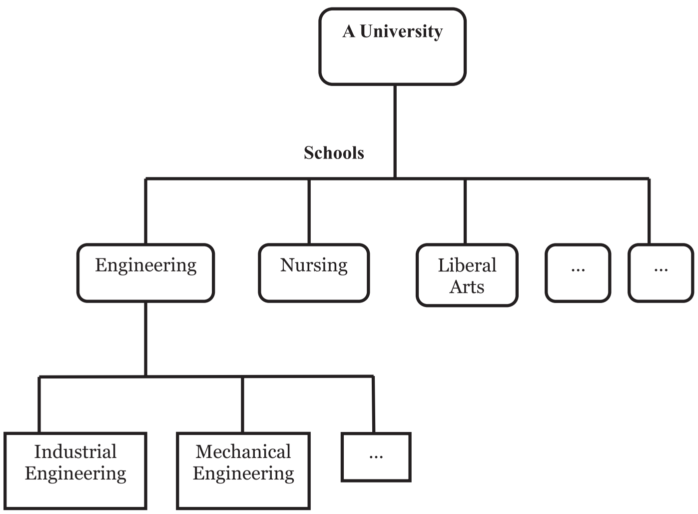
RESEARCH METHODOLOGIES

Different forecasting models are developed to forecast the number of undergraduate student enrollment; we exclude first-semester freshman students in this research. Based on a detailed comparison and the analysis of various forecasting models, this research endeavor is expected to identify the most suitable forecasting model for the projection of undergraduate student enrollment at a university. We cannot assume this method will be best for all institutions, but put it forth for consideration by the reader. It is important to highlight the principal model-related assumptions made in this research:

1. Student enrollment patterns that have happened in the past are considered likely to occur in the future for enrollment forecasting purposes.
2. Available historical data for analysis can be assumed to represent the entire historical pattern that can be used to predict future student enrollment.
3. Student return ratios are different from fall to spring semesters.

Forecasting undergraduate student enrollment is primarily based on historical data. Three levels of analysis—university, school, and division levels—are analyzed to forecast undergraduate student enrollment each semester, including university analyses. In this research, a university is assumed to consist of several schools (e.g., engineering or nursing), each having multiple divisions (e.g., industrial engineering or mechanical engineering) as shown in Figure 1.

Figure 1. Overview of a University, Schools, and Divisions



University-level analysis forecasts student enrollment based on return ratio calculations of total student enrollment each semester, which is similar to many proposed methods in the literature that have shown high accuracy in projecting student enrollment as discussed in the previous section of the paper. School- and division-level analyses, on the other hand, have not been mentioned in the literature we reviewed since these are the two new units of analysis proposed in this research. School-level analysis forecasts student enrollment based on return ratio calculations of each school's student enrollment to calculate total student enrollment each semester. Division-level analysis forecasts student enrollment, based on return ratio calculations of each division's student enrollment, to calculate total student enrollment each semester.

All three forecasting models used two types of return ratio (RR) calculations: average return ratio (ARR) and exponential weighted return ratio (EWRR). We based all calculations for forecasting a campus's student enrollment on calculating the ratio between student enrollment in a specific semester and the previous semester. The mathematical equation of each RR can be expressed as

$$RR_j = \frac{E_t}{E_{t-1}} \tag{1}$$

where E_t is the undergraduate student enrollment at semester t.

University-Level Analysis

University-level analysis forecasts undergraduate student enrollment, based on ARR and EWRR calculations of a university's student enrollment each semester from the previous semester. When calculating ARR and EWRR, fall semesters and spring semesters are separated to increase the forecasting accuracy. The ARR of university-level analysis is expressed as

$$ARR_t = \frac{1}{n}\sum_{t=1}^{n} RR_t \tag{2}$$

where n is the total number of semesters. The EWRR of university-level analysis can also be calculated by

$$EWRR_t = \alpha[RR_{t-1} + (1-\alpha)RR_{t-2} + (1-\alpha)^2 RR_{t-3} + \ldots] + (1-\alpha)^{t-1} RR_1 \tag{3}$$

where α is a weighting factor ($0 < \alpha < 1$); as α increases, more weight is given to recent data. Based on Equations (1) and (2), we can obtain undergraduate student enrollment projections, using university-level analysis, by

$$E_t = ARR_t \times E_{t-1} \tag{4}$$

$$E_t = EWRR_t \times E_{t-1} \tag{5}$$

School-Level Analysis

It is possible that university-level analysis may not capture different students' retention rates in different schools. As a result, we propose school-level forecasting analysis to (potentially) increase the forecasting accuracy. In this analysis, we project undergraduate student enrollment, based on each school's student enrollment, and calculate the total student enrollment, considering the differences between the fall and spring semesters' student RR. The ARR of school-level analysis is calculated by

$$ARR_{t,j} = \frac{1}{n}\sum_{t=1}^{n} RR_t \quad (6)$$

where $E_{t,j}$ represents the undergraduate student enrollment in school j at semester t. The EWRR for the school-level analysis is calculated by

$$EWRR_{t,j} = \alpha[RR_{t-1} + (1-\alpha)RR_{t-2} + (1-\alpha)^2 RR_{t-3} + \ldots] + (1-\alpha)^{t-1} RR_1 . \quad (7)$$

Therefore, the total student enrollment, using school-level analysis, can be obtained by

$$E_{t,j} = ARR_{t,j} \times E_{t-1,j} \quad (8)$$

$$E_{t,j} = EWRR_{t,j} \times E_{t-1,j} . \quad (9)$$

Division-Level Analysis

In this approach, different schools are broken down into different divisions to see if forecasting accuracy can be further improved. After we obtain student enrollment from each division, we can calculate the total student enrollment. When calculating ARR and EWRR, we also separate fall semesters and spring semesters. The ARR for the division-level analysis is calculated by

$$ARR_{t,k} = \frac{1}{n}\sum_{t=1}^{n} RR_t \quad (10)$$

where $E_{t,k}$ is the undergraduate student enrollment in division k at semester t. The EWRR for division-level analysis is calculated by

$$EWRR_{t,k} = \alpha[RR_{t-1} + (1-\alpha)RR_{t-2} + (1-\alpha)^2 RR_{t-3} + \ldots] + (1-\alpha)^{t-1} RR_1 . \quad (11)$$

Then, the total student enrollment can be obtained as follows:

$$E_{t,k} = ARR_{t,k} \times E_{t-1,k} \quad (12)$$

$$E_{t,k} = EWRR_{t,k} \times E_{t-1,k} . \quad (13)$$

In this study, the forecasting errors are measured to compare the performance of different forecasting models, and the forecasting error, ε (%), is defined as

$$\varepsilon = \frac{E^A_{total} - E^P_{total}}{E^A_{total}} \times 100 \quad (14)$$

where E^A_{total} and E^P_{total} are the total numbers of actual and projected student enrollment, respectively. An example of forecasting analysis is given below.

Numerical Results: Forecasting Model Illustration

Table 1 shows sample data from the fall and spring semesters of 2000 to 2010 at the State University of New York (SUNY) at Binghamton, where Y_{00f} represents the fall semester of 2000, Y_{01s} represents the spring semester of 2001, s_1 represents first semester, s_2 represents second semester, and so on. Undergraduate students typically take 4 years (or eight semesters) to graduate; there are some students, however, who take more than 4 years, which is why the data include student enrollment for up to 6 years.

Table 1. Sample Data of Undergraduate Student Enrollment (2000–2010)

Academic Year	Semester					
	s_1	s_2	…	s_{11}	s_{12}	Total
Y_{00f}	879					
Y_{01s}	95	830				
⋮			⋮			⋮
Y_{10s}	135	900	...	7	8	3,690
Y_{10f}	1,089	124	...	9	2	3,796

To calculate the forecasting error, using the university-level analysis, the RR should be calculated by Equation (1), as shown in Table 2. For instance, the RR of first-semester students becoming second-semester students from the fall semester of 2000 to the spring semester of 2001 is calculated as

$$RR_{Y_{00f},Y_{01s}}(s_{1,2}) = \frac{E_2(Y_{01s})}{E_1(Y_{00f})} = \frac{830}{879} = 0.94. \quad (15)$$

The RR of the 11th-semester students becoming 12th-semester students from the spring semester of 2010 to the fall semester of 2010 is calculated as

$$RR_{Y_{108s},Y_{10f}}(s_{11,12}) = \frac{E_8(Y_{10f})}{E_7(Y_{10s})} = \frac{2}{7} = 0.29. \quad (16)$$

Table 2. Summary of RRs for Each Semester

Academic Year	Semester				
	s_1	s_2	…	s_{11}	s_{12}
$Y_{00f,01s}$	0.94				
$Y_{01s,01f}$	0.85	0.91			
⋮			⋮		
$Y_{08f,09s}$	0.96	1.03	...	1.29	0.42
$Y_{09s,09f}$	0.98	0.98	...	0.22	0.56

After all the RRs are calculated, we can apply the two forecasting methods, ARR and EWRR. The ARRs between each semester are calculated by Equation (2) as shown in Table 3, where $ARR_{F,S}$ represents the ARR from fall to spring semesters and $ARR_{S,F}$ represents the ARR from spring to fall semesters. For instance, the $ARR_{F,S}(s_{1,2})$ is equal to the average of all the RRs from fall to spring semesters in the $s_{1,2}$ column in Table 2.

$$ARR_{F,S}(s_{1,2}) = \frac{1}{9}(0.94 + 0.95 + ... + 0.96) = 0.96. \quad (17)$$

The $ARR_{S,F}(s_{5,6})$ is equal to the average of all the RRs from spring to fall semesters between the fifth and the sixth semesters.

$$ARR_{S,F}(s_{5,6}) = \frac{1}{7}(0.67 + 0.91 + ... + 0.89) = 0.83. \quad (18)$$

Based on ARRs, student enrollment can be projected for Y_{10s} and Y_{10f} using Equation (4), as shown in Table 4. When projecting for Y_{10s} , we use the $ARR_{F,S}$ since it projects for a spring semester. When projecting for Y_{10f} , however, we use the $ARR_{S,F}$ since it projects for a fall semester. For instance, the 2nd- and 10th-semesters' student enrollment at Y_{10s} and Y_{10f} can be expressed as

$$E_{10s}(s_2) = E_{09f}(s_1) \times ARR_{F,S}(s_{1,2}) = 930 \times 0.96 = 892.8 \quad (19)$$

$$E_{10f}(s_{10}) = E_{10s}(s_9) \times ARR_{S,F}(s_{9,10}) = 21 \times 0.32 = 7. \quad (20)$$

We calculate the forecasting errors projecting for the spring and fall semesters of 2010, using university-level analysis with the ARR, by Equation (14) as shown in Table 5. The detailed calculations are illustrated as

$$\varepsilon(Y_{10s}) = \frac{E^A_{total}\ (Y_{10s}) - E^P_{total}\ (Y_{10s})}{E^A_{total}\ (Y_{10s})} \times 100 = \frac{3555 - 3532}{3555} \times 100 = 0.66\% \quad (21)$$

$$\varepsilon(Y_{10f}) = \frac{E^A_{total}\ (Y_{10f}) - E^P_{total}\ (Y_{10f})}{E^A_{total}\ (Y_{10f})} \times 100 = \frac{2707 - 2718}{2707} \times 100 = -0.37\%. \quad (22)$$

Table 5. Summary of ε for Y_{10s} and Y_{10f} Using ARR

	E^A_{total}	E^P_{total}	ε
Y_{10s}	3555	3532	0.66
Y_{10f}	2707	2718	–0.37

Table 3. Summary of ARRs for Each Semester

	$s_{1,2}$	$s_{2,3}$	$s_{3,4}$	$s_{4,5}$	$s_{5,6}$	$s_{6,7}$	$s_{7,8}$	$s_{8,9}$	$s_{9,10}$	$s_{10,11}$	$s_{11,12}$
$ARR_{F,S}$	0.96	0.99	0.96	0.93	0.94	0.97	0.93	0.83	0.63	0.94	0.61
$ARR_{S,F}$	0.90	0.91	0.91	0.87	0.83	0.85	0.36	0.16	0.32	0.20	0.45

Table 4. Summary of $E^P(Y_{10s})$ and $E^P(Y_{10f})$ Using ARR

	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}
$E^P(Y_{10s})$	890	107	905	126	730	109	580	21	51	6	7
$E^P(Y_{10f})$	121	819	104	792	111	623	37	91	7	10	3

On the other hand, we calculate the EWRR between each semester using Equation (3) based on the results in Table 2. The calculated EWRRs are shown in Table 6, where $EWRR_{F,S}$ represents the EWRR from fall to spring semesters, and the $EWRR_{S,F}$ represents the EWRR from spring to fall semesters. Assuming that α is set to 0.5 for the purpose of this research, then we can calculate $EWRR_{F,S}$ and $EWRR_{S,F}$ as

$$EWRR_{F,S}(s_{3,4}) = 0.5\times[0.99+(1-0.5)0.96+(1-0.5)^2 0.95+(1-0.5)^3 0.95+ (1-0.5)^4 0.96+(1-0.5)^5 0.96+(1-0.5)^6 0.95]+(1-0.5)^7 0.96 \quad (23)$$

$$EWRR_{S,F}(s_{9,10}) = 0.5\times[0.34+(1-0.5)0.24+(1-0.5)^2 0.33+(1-0.5)^3 0.36+ (1-0.5)^4 0.34 = 0.31. \quad (24)$$

We can now project student enrollment using the EWRR for Y_{10s} and Y_{10f} using Equation (5), as shown in Table 7. When projecting for Y_{10s}, we use the $EWRR_{F,S}$ since it projects for a spring semester. When projecting for Y_{10f}, however, we use the $EWRR_{S,F}$ since it projects for a fall semester, as illustrated in the following examples:

$$E_{10s}(s_3) = E_{09f}(s_2)\times EWRR_{F,S}(s_{2,3}) = 108\times 1.01 = 109 \quad (25)$$

$$E_{10f}(s_8) = E_{10s}(s_7)\times EWRR_{S,F}(s_{7,8}) = 103\times 0.31 = 32\ . \quad (26)$$

We also can calculate the forecasting errors, using university-level analysis with the EWRR, as shown in Table 8. The detailed calculations are illustrated as

$$\varepsilon(Y_{10s}) = \frac{E^A_{total}\ (Y_{10s}) - E^P_{total}\ (Y_{10s})}{E^A_{total}\ (Y_{10s})}\times 100 = \frac{3555-3543}{3555}\times 100 = 0.35\% \quad (27)$$

$$\varepsilon(Y_{10f}) = \frac{E^A_{total}\ (Y_{10f}) - E^P_{total}\ (Y_{10f})}{E^A_{total}\ (Y_{10f})}\times 100 = \frac{2707-2724}{2707}\times 100 = -0.69\%\ . \quad (28)$$

Table 8. Summary of ε for Y_{10s} and Y_{10f} Using EWRR

	E^A_t	E^P_t	ε
Y_{10s}	3555	3543	0.35
Y_{10f}	2707	2724	–0.69

We illustrate the calculation steps through an example using university-level analysis, which forecasts undergraduate student enrollment based on the ARR and EWRR calculations of a university's student enrollment each semester from the previous semester. School- and division-level analyses, on the other hand, forecast undergraduate student enrollment, based on the ARR and EWRR calculations of different schools' or different divisions' student enrollment. The detailed calculation steps are very similar for school- and division-level analyses compared to university-level analysis. The only difference is that after we obtain student enrollment at each school or each division, we calculate the total student enrollment at a university by taking the summation of all schools' or divisions' student enrollment, after which we compare the resulting E^P_{total} with E^A_{total} to obtain ε.

To find the most accurate forecasting model to project undergraduate student enrollment at a university, we compare the forecasting error, ε, for the three forecasting models—university-, school-, and division-level analyses—and the forecasting model with the lowest ε value should be selected. As shown in Table 9, we compare ε value of each forecasting model combined with its corresponding method. It is evident that university- and school-level analyses yield lower ε values than division-level analysis when forecasting for undergraduate student enrollment (excluding first-semester freshman students) for the spring and fall semesters of 2010.

Table 6. Summary of EWRRs for Each Semester

	$s_{1,2}$	$s_{2,3}$	$s_{3,4}$	$s_{4,5}$	$s_{5,6}$	$s_{6,7}$	$s_{7,8}$	$s_{8,9}$	$s_{9,10}$	$s_{10,11}$	$s_{11,12}$
$EWRR_{F,S}$	0.96	1.01	0.96	0.94	0.94	0.96	0.93	0.84	0.64	1.06	0.57
$EWRR_{S,F}$	0.95	0.91	0.95	0.88	0.87	0.85	0.31	0.16	0.31	0.21	0.45

Table 7. Summary of $E^P(Y_{10s})$ and $E^P(Y_{10f})$ Using EWRR

	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}
$E^P(Y_{10s})$	894	109	910	126	730	108	579	21	52	7	7
$E^P(Y_{10f})$	128	816	108	793	116	620	32	91	7	10	3

Table 9. Comparison of ε Values of Y_{10s} and Y_{10f}

	University-Level		School-Level		Division-Level	
	ARR	EWRR	ARR	EWRR	ARR	EWRR
Y_{10s}	0.66	0.35	0.46	0.38	–1.94	–0.87
Y_{10f}	–0.38	–0.69	–0.83	–1.13	–3.44	–1.10

Table 10. Comparison of ε Values of Y_{10s} to Y_{10f}

	University-Level		School-Level	
	ARR	EWRR	ARR	EWRR
Y_{07s}	–0.82	–0.76	–1.17	–1.02
Y_{07f}	–1.64	–1.67	–1.97	–1.70
Y_{08s}	0.23	0.55	–0.13	0.33
Y_{08f}	–0.33	0.12	–0.86	–0.22
Y_{09s}	1.02	1.09	0.60	0.84
Y_{09f}	1.25	1.60	0.93	1.48

To determine whether university- and school-level analyses will work in forecasting undergraduate student enrollment during special periods or sudden events (e.g., the economic downturn in 2008), we also analyze the forecasting results from the spring semester of 2007 to the fall semester of 2009 to validate their ε values, as shown in Table 10. It is clear that the forecasting models also work relatively well in forecasting undergraduate student enrollment during such events.

In this research, the forecasting models are mainly developed by the ARR and EWRR based on all available data prior to the projected semester. To further determine whether university- or school-level analysis is better, we will use only the most recent data points for analysis. We compare different ε values based on 1 year (y^1), 2 years (y^2), 3 years (y^3), and 4 years (y^4) of the ARRs, which means that we calculate the ARRs based on using only the most recent years of enrollment data for calculations. Since the calculations are now based on a small numbers of years, the EWRR method will not differ significantly from the ARR; therefore, we will use only the ARR for the analysis. By using the same calculation steps, we show the results for ε values using university-level analysis in Table 11 and Figure 2.

Table 11. ε Values When Using Different Years of ARR by University-Level Analysis for Projecting Y_{10s} and Y_{10f}

	y_1	y_2	y_3	y_4
Y_{10s}	–0.20	0.23	0.68	0.76
Y_{10f}	–1.55	–0.69	–0.12	–0.29

Figure 2. ε Values When Using Different Years of ARR by University-Level Analysis for Projecting Y_{10s} and Y_{10f}

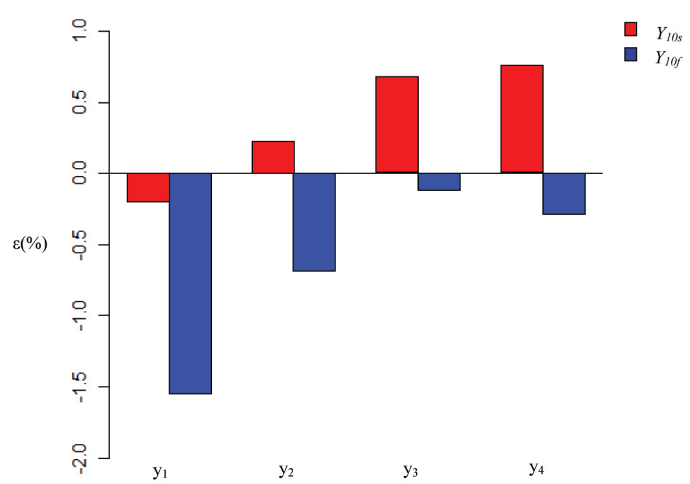


Table 12. ε **Values of Using Different Years of ARR by School-Level Analysis for Projecting** Y_{10s} **and** Y_{10f}

	y_1	y_2	y_3	y_4
Y_{10s}	–0.08	0.29	0.76	0.79
Y_{10f}	–1.67	–1.08	–0.49	–0.55

Figure 3. ε **Values When Using Different Years of ARR by School-Level Analysis for Projecting** Y_{10s} **and** Y_{10f}

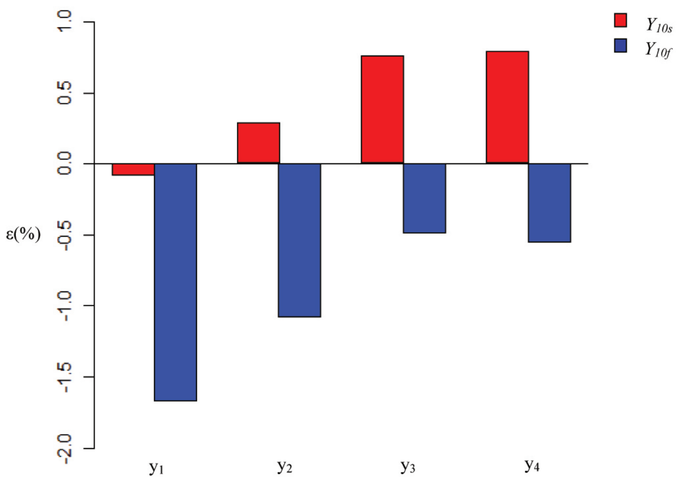


Figure 4. ε **Value Comparison for Projecting** Y_{10s} **and** Y_{10f}

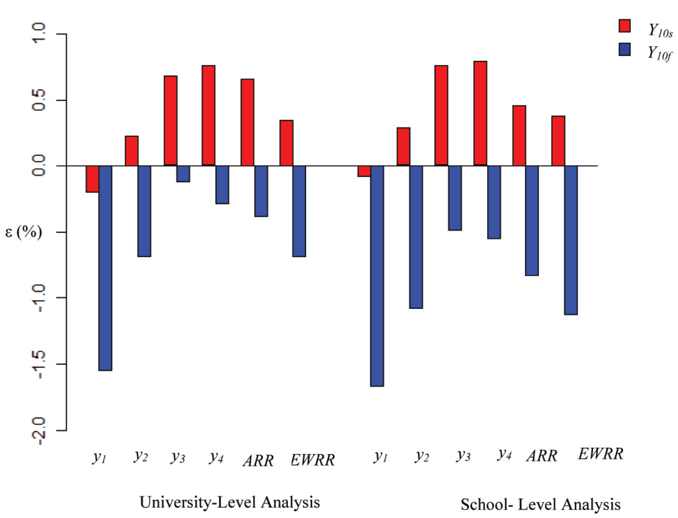


Table 13. ε **Value Comparison Using ARR and EWRR for Projecting** Y_{09s} **and** Y_{09f}

	University-Level		School-Level	
	ARR	EWRR	ARR	EWRR
Y_{09s}	1.02	1.09	0.60	0.84
Y_{09f}	1.25	1.60	0.93	1.49

By using the same calculation steps, we show the results for ε values using school-level analysis in Table 12 and Figure 3. The results indicate that when the average number of years used increases, $\varepsilon(Y_{10s})$ generally increases as well. However, $\varepsilon(Y_{10f})$ generally decreases from 1 year to 3 years, with a slight increase from 3 years to 4 years.

To illustrate the lowest ε value, we develop Figure 4 for comparison purposes. It is evident that when projecting for Y_{10s}, school-level analysis using 1-year average provides the lowest ε value. When projecting for Y_{10f}, university-level analysis using 3 years of average yields the lowest ε value.

To examine if the forecasting models can be extended to project for spring and fall semesters, we performed the same calculation steps for forecasting undergraduate student enrollment at a university, projecting for Y_{09s} and Y_{09f}. We summarize $\varepsilon(Y_{09s})$ and $\varepsilon(Y_{09f})$ for both university- and school-level analyses in Table 13.

By using the same calculation steps, we show ε values using university-level analysis for projecting Y_{09s} and Y_{09f} in Table 14 and Figure 5.

Table 14. ε **Values When Using Different Years of ARR by University-Level Analysis for Projecting** Y_{09s} **and** Y_{09f}

	y_1	y_2	y_3	y_4
Y_{09s}	0.85	1.29	1.27	1.16
Y_{09f}	1.58	2.00	1.57	1.58

Figure 5. ε Values When Using Different Years of ARR by University-Level Analysis for Projecting Y_{09s} and Y_{09f}

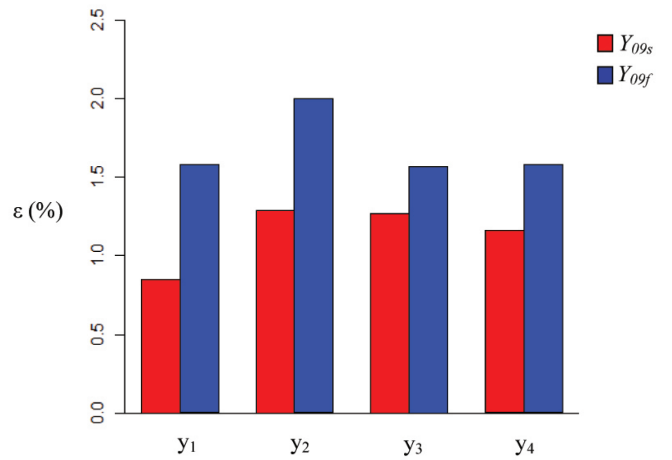


Table 15. ε Values When Using Different Years of ARR by School-Level Analysis for Projecting Y_{09s} and Y_{09f}

	y_1	y_2	y_3	y_4
Y_{09s}	0.71	1.11	1.01	0.68
Y_{09f}	1.57	1.85	1.46	1.70

By using the same calculation steps, we show ε values using school-level analysis for projecting Y_{09s} and Y_{09f} in Table 15 and Figure 6.

When finding the lowest ε value for projecting Y_{09s} and Y_{09f} we plotted Figure 7 for comparison purposes. Based on this comparison, it is clear that when projecting for Y_{09s}, school-level analysis based on the ARR yields the lowest ε value. When projecting for Y_{09f}, school-level analysis based on the ARR also provides the lowest ε. Hence, by combining the two cases, we recommend school-level analysis based on a 1-year average (i.e., lowest forecasting error) when projecting for spring semesters. On the other hand, we recommend university-level analysis using the ARR method since this forecasting model yields the lowest forecasting error when projecting for fall semesters. Based on the results, it is reasonable that averaging over more years of data provides a better forecast result when projecting for fall semesters since the data are noisy. This is mainly due to students taking internships or transferring to and from another school starting at the fall semester. On the other hand, forecasting for spring semester using 1 year of data is sufficient since the data are more stable.

Figure 6. ε Values When Using Different Years of ARR by School-Level Analysis for Projecting Y_{09s} and Y_{09f}

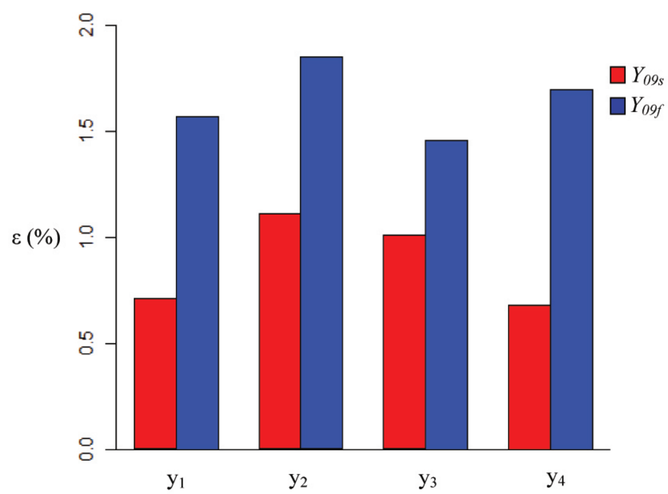
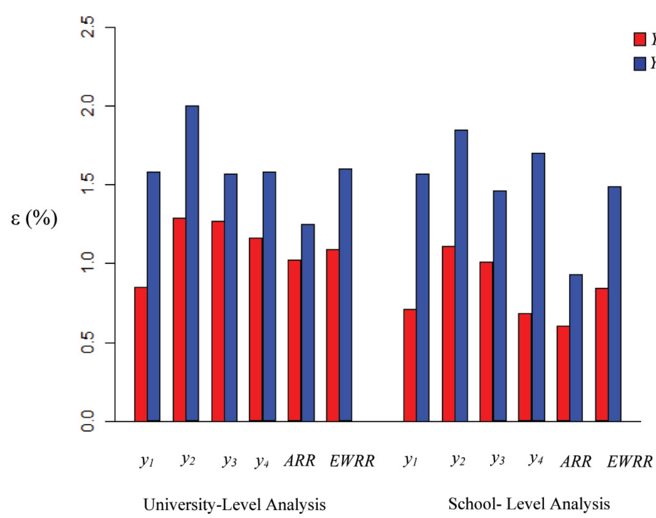


Figure 7. ε Value Comparison for Projecting Y_{09s} and Y_{09f}



CONCLUSION AND FUTURE WORK

This research focuses on developing a low-cost and easy-to-use forecasting model for projecting undergraduate student enrollment. Based on a detailed analysis of historical data and different forecasting models, we developed and evaluated two forecasting models using different sets of enrollment data, including university-, school-, and division-level enrollment. University-level analysis is similar to many proposed methods in the literature that have shown high accuracy in projecting student enrollment as discussed in the previous sections of the paper. School- and division-level analyses were not mentioned in the literature we reviewed since these are the units of analysis to test the two new methods proposed in this research. The numerical results indicate that the forecasting errors will not decrease when applying division-level analysis versus school- and university-level analyses. Also, using all available student enrollment data does not necessarily produce a smaller forecasting error than using the most recent enrollment data. Based on the case study, by looking at different years' forecasting errors, school-level analysis using 1-year average should be used when projecting for spring semesters since the model yields the lowest average forecasting error of 0.40%. When projecting for fall semesters, university-level analysis using the ARR method should be used since it yields the lowest average forecasting error of 0.81%. Therefore, to keep the forecasting error rate at the lowest level, it is better to use school-level analysis with 1-year average when projecting for spring semesters and to use university-level analysis with the ARR when projecting for fall semesters.

The research is based on analyzing historical undergraduate student enrollment data from the State University of New York at Binghamton by comparing forecasting errors of different forecasting models. The proposed forecasting models should be updated constantly with current and accurate information regarding student enrollment data, such that new enrollment trends can be analyzed. A user-friendly graphical user interface can also be implemented and applied in the future to make the computations of the forecasting models more efficient and effective.

References

Armstrong, J. S. 2001. *Principles of forecasting: A handbook for researchers and practitioners.* Norwell, MA: Kluwer Academic.

Beck, B. D. 2009. Enrollment forecasting and modeling. Office of Academic Planning and Analysis, University of Wisconsin, Madison.

Brown, R. G. 1959. S*tatistical forecasting for inventory control.* New York: McGraw-Hill.

Desjardins, S. L., Ahlburg, D. A., & McCall, B. P. 2006. An integrated model of application, admission, enrollment, and financial aid. *Journal of Higher Education, 77*(3), 381–429.

Dobbs, S. 2001. Re-engineering the enrollment management system at the Monterey Peninsula Unified School District (MPUSD). MS thesis, Naval Postgraduate School, Monterey, CA.

Gardner, D. E. 1981. Weight factor selection in double exponential smoothing enrollment forecasts. *Research in Higher Education, 14*(1), 49–56.

Glover, R. H. 1986. Designing a decision-support system for enrollment management. *Research in Higher Education, 24*(1), 15–34.

Guo, S. 2002. Three enrollment forecasting models: Issues in enrollment projections for community colleges. Fortieth RP Conference, Pacific Grove, CA, May 1–3.

Holt, C. C. 2004. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting, 20*(1), 5–10.

Hossler, D., & Bean, J. P. 1990. *The strategic management of college enrollments.* San Francisco: Jossey-Bass.

Magee, J. F. 1958. *Production planning and inventory control.* New York: McGraw-Hill.

Marshall, K. T., & Oliver, R. M. 1970. A constant-work model for student attendance and enrollment. *Operations Research, 18*(2), 193–206.

Muth, J. F. 1960. Optimal properties of exponentially weighted forecasts. *Journal of the American Statistical Association, 55*(290), 299–306.

Norton, D. R. 1998. Universities' enrollment management. State of Arizona Office of the Auditor, Phoenix, AZ.

Porter, S. R. 1999. Including transfer-out behavior in retention models: Using the NSLC enrollment search data. North East Association of Institutional Research Conference, Newport, RI.

Schmid, C. F., & Shanley, F. J. 1952. Techniques of forecasting university enrollment. *Journal of Higher Education, 23*(9), 483–488, 502–503.

Snyder, R. D. 1988. Progressive tuning of simple exponential smoothing forecasts. *Journal of the Operational Research Society, 39*(4), 393–399.

Tersine, R. J. 1994. *Principles of inventory and materials management*. Boston: PTR Prentice-Hall Publishing.

Appendix. List of Nomenclature

t	A specific semester
j	A specific school
k	A specific division
n	Total number of semesters
s_t	Semester t
$s_{\beta,\theta}$	Two consecutive semesters, where $\beta < \theta$
y	Number of years used to calculate ARR
ARR_t	Average return ratio for semester t
$ARR_{t,j}$	Average return ratio for semester t and school j
$ARR_{t,k}$	Average return ratio for semester t and division k
$ARR_{F,S}$	Average return ratio from fall to spring semester
$ARR_{S,F}$	Average return ratio from spring to fall semester
E	Undergraduate student enrollment
E_t	Number of undergraduate student enrollment at semester t
$E_{t,j}$	Number of undergraduate student enrollment at semester t in school j
$E_{t,k}$	Number of undergraduate student enrollment at semester t in division k
E_{t-1}	Number of undergraduate student enrollment at semester $t-1$
$E_{t-1,j}$	Number of undergraduate student enrollment at semester $t-1$ in school j
$E_{t-1,k}$	Number of undergraduate student enrollment at semester $t-1$ in division k
E_{total}	Total undergraduate student enrollment
E_{total}^{A}	Actual total undergraduate students
E_{total}^{P}	Projected total undergraduate students
$EWRR_t$	Exponential weighted return ratio for semester t
$EWRR_{t,j}$	Exponential weighted return ratio for semester t and school j
$EWRR_{t,k}$	Exponential weighted return ratio for semester t and division k
$EWRR_{F,S}$	Exponential weighted ratio from fall to spring semester
$EWRR_{S,F}$	Exponential weighted return ratio from spring to fall semester
RR	Return ratio
RR_t	Return ratio for semester t
RR_{t-1}	Return ratio for semester $t-1$
Y	Academic year
Y_f	Fall academic year
Y_s	Spring academic year
α	Weighting factor for EWRR method, where $0 < \alpha < 1$
ε	Forecasting error (%)