



SCHOOL of
GRADUATE STUDIES
EAST TENNESSEE STATE UNIVERSITY

East Tennessee State University
Digital Commons @ East
Tennessee State University

Electronic Theses and Dissertations

Student Works


12-2017

Modeling Student Enrollment at ETSU Using a Discrete-Time Markov Chain Model

Lohuwa Mamudu

East Tennessee State University

Follow this and additional works at: <https://dc.etsu.edu/etd>

 Part of the [Adult and Continuing Education Administration Commons](#), and the [Science and Mathematics Education Commons](#)

Recommended Citation

Mamudu, Lohuwa, "Modeling Student Enrollment at ETSU Using a Discrete-Time Markov Chain Model" (2017). *Electronic Theses and Dissertations*. Paper 3310. <https://dc.etsu.edu/etd/3310>

This Thesis - Open Access is brought to you for free and open access by the Student Works at Digital Commons @ East Tennessee State University. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ East Tennessee State University. For more information, please contact digilib@etsu.edu.

Modeling Student Enrollment at East Tennessee State University Using a
Discrete-Time Markov Chain Model

A thesis

presented to

the faculty of the Department of Mathematics

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Master of Science in Mathematical Sciences

by

Lohuwa Mamudu

December, 2017

Michele Joyner Ph.D., Chair

Christina N. Lewis Ph.D.

Robert Gardner Ph.D.

Keywords: Discrete-Time Markov Chain Model

ABSTRACT

Modeling Student Enrollment at East Tennessee State University Using a Discrete-Time Markov Chain Model

by

Lohuwa Mamudu

Discrete-time Markov chain models can be used to make future predictions in many important fields including education. Government and educational institutions today are concerned about college enrollment and what impacts the number of students enrolling. One challenge is how to make an accurate prediction about student enrollment so institutions can plan appropriately. In this thesis, we model student enrollment at East Tennessee State University (ETSU) with a discrete-time Markov chain model developed using ETSU student data from Fall 2008 to Spring 2017. In this thesis, we focus on the progression from one level to another within the university system including graduation and dropout probabilities as indicated by the data. We further include the probability that a student will leave school for a limited period of time and then return to the institution. We conclude with a simulation of the model and a comparison to the trends seen in the data.

Copyright by Lohuwa Mamudu 2017

All rights reserved.

ACKNOWLEDGEMENTS

First and Foremost, I thank the Almighty Allah for how far He has brought me and seeing me through this thesis. I would like to acknowledge Dr. Michele Joyner for her assistance, sense of direction and support throughout this thesis. When everything seems tougher, she helped me rise above all the difficult moments in the course of this thesis. She was a mother and not just my adviser. I am very grateful and thankful to her for accepting me as one of her thesis student. It was a pleasure an a great privilege working with her.

I would also like to thank Dr Nicole Lewis and Dr Cintron-Arias Ariel for agreeing to be part of my committee. It is out of their immense knowledge and support that help brought this thesis to fruition.

Also, a special thanks to my brother Dr Hadii Mamudu, professor at the Department of Public Health, ETSU. It is mostly through his productive advice and encouragement, streamlined me in the academic field.

Finally, I acknowledge and thank my family, friends and loved ones for their lasting support, love and prayers.

TABLE OF CONTENTS

ABSTRACT	2
ACKNOWLEDGEMENTS	4
LIST OF TABLES	7
LIST OF FIGURES	8
1 INTRODUCTION	9
1.1 Background	10
1.2 Overview of Thesis	11
2 DESCRIPTION AND PRELIMINARY ANALYSIS OF THE DATA SET	13
2.1 Data Description	13
2.2 Trend in Student Enrollment	15
3 MODEL DEVELOPMENT AND ANALYSIS	21
3.1 Discrete-Time Markov Chain Probability Transition Diagram	21
3.2 Discrete-Time Markov Chain Transition Probability State Matrices	23
3.2.1 Matrix Comparison	28
4 DISCUSSION AND ANALYSIS OF THE HIATUS STATE	32
4.1 Development of the Hiatus Model	40
4.2 Analysis and Long-Term Future Prediction	46
5 MODELING THE ENROLLMENT OF NEW STUDENTS	51
6 SIMULATION OF THE MODEL	54
7 CONCLUSION AND FUTURE WORK	59

REFERENCES 61
VITA 65

LIST OF TABLES

1	Description of Factors in the Data Set	14
2	A table showing the description of variables to analyze for students taking a break	33
3	Prediction of Long-Term Probability Transitions of the States	49

LIST OF FIGURES

1	Trend of undergraduates student enrollment.	15
2	Trend of undergraduate student enrollment by class level.	17
3	Trend of senior graduation.	17
4	Trend of undergraduate student dropout.	19
5	Trend of undergraduate student dropout for each class levels.	19
6	Discrete-Time Markov Chain Probability Transition Diagram	22
7	Number of students who transitions to the various classes after hiatus	35
8	Side by side boxplot on the number of semester spent on hiatus . . .	36
9	Bar plot of student status left	37
10	Bar plot of student status Return	38
11	Bar plot of Freshman status Return	39
12	Bar plot of Sophomore status Return	39
13	Bar plot of Junior status Return	40
14	The transition probability diagram for hiatus.	41
15	The Total New Student Enrolled	51
16	The Total New Student Enrolled By Class Level	53
17	Simulation of Total Student Enrolled	56
18	Simulation of Total Student Enrolled By Class Level	56
19	Simulation of Total New Student Enrolled	57
20	Simulation of Total New Student Enrolled by Class Level	57

1 INTRODUCTION

The enrollment of students in a university system across the nation between 2001 and 2011 increased by 32%, from 15.9 million to 21.0 million [1]. In 2014, reported federal projections showed college enrollment would rise by only 14% through 2022 which is about a third of the pace in the past decade. The enrollment figures were contained in a broader report released by the National Center for Education Statistics[1].

For years, America's college campuses swelled with more and more students, but enrollment peaked in 2010 at just over 21 million students. Attendance has dropped every year since 2010 [2]. There were over 800,000 fewer students walking around college campuses by fall 2014 [2]. Some indicate that there is no need to be remorse; the drop is happening as a result of improvement in the economy. Their rationale is that more people are going back to work instead of signing up for additional degrees [2]. President Obama, in his statement last year said, "A college degree is the surest ticket to the middle class" [2]. He proposed making community colleges free for two years, because he perceived them as the gateway to higher education and better jobs for many Americans [2].

Having a college degree versus only a high school diploma makes a big difference when it comes to employment and wages [3]. About only 2.5% of college graduates are unemployed, compared to 5.6% of high school graduates [3]. In 2015, college enrollment in the spring semester fell nearly 2% from the previous year, to 18.6 million in the U.S. [4]. The bulk of the decline came among students over 24, whose numbers fell by 264,000, or 3.6 percent [4]. Among the 1,271 ranked schools that submitted this data to U.S. News in an annual survey, the average undergraduate enrollment

for fall 2015 was slightly over 6,000 [5]. The average among the 10 schools with the highest undergraduates enrollment was more than seven times higher at 43,936 [5]. These statistics indicates there is a continuous fall in college enrollment in the U.S.

1.1 Background

Discrete-time Markov chains have been used over some years to date in several disciplines including health in predicting the disease progression [8], education in predicting enrollment [10], and other forecast projects based on present events. For instance, a discrete-time Markov chain model has been used to forecast daily admission scheduling and resource planning in a cost or capacity constrained health care system [6]. Another area where discrete-time Markov chain have been used was in investigating the effects of treatment programs and health care protocols for chronic diseases [8]. One major characterization of discrete-time Markov chains is that the next event to occur depends only on the present state of the system and does not depend on the history of the system. Hence, this Markov chain process is said to be “memory-less”.

A discrete-time Markov chain is said to be a stochastic process which satisfies the Markov property given by

$$P(X_{n+1} = j | X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i) \quad (1)$$

where X_n is a sequence of random variables with discrete time steps t_0, t_1, \dots, t_n . A discrete-time Markov chain can be homogeneous or non-homogeneous. A homogeneous Markov chain is the one in which the probability transition matrices do not change over time from one state to the other (i.e cycles) and across subjects. In a

non-homogeneous Markov chain, the probability transition matrices change over time from state to state and/or across subjects [19].

In this thesis, we use a discrete-time Markov chain model to predict student enrollment at ETSU. We assume that the discrete-time Markov chain is non-homogeneous so that the probability transition matrix changes over the observation time. A discrete-time Markov chain model is applicable to this study, because the state to which a student transitions depends on only the current state (or academic year) and not on the past. Therefore, student enrollment from semester to semester satisfies the Markov property with the time increments being semesters. This makes it possible for us to model student enrollment and then use the model to predict graduation rates, dropout rates and total student enrollment using data obtained on student progression. Enrollment retention rates are an accepted indicator of a university's success in providing quality degree programs and learning environments which will lead to a student's continued enrollment and timely graduation from the university campus [9].

1.2 Overview of Thesis

We will be using a discrete-time Markov chain to predict student enrollment based on transitioning from a freshman to a senior. Similar work was done by Helbert [10] in 2015 which focused on modeling enrollment at a regional university using a discrete-time Markov chain. The states considered in his model included freshman, sophomore, junior, senior, drop out, and graduate, similar to our model as discussed below. However, his model predicted that 21% of students graduated in five years

[10]. This was seen as a discrepancy in his model prediction because his data initially indicated that 1,085 students graduated in five years out of 3,079 total freshmen. This represents 35% of freshman present in Fall 2008 graduated by Spring 2013 (5 years) [10]. In this thesis, we identify some of the discrepancies and try to correct them to enhance our model prediction.

This thesis is organized into seven chapters. The first chapter introduces and gives some background information about student enrollment and discrete-time Markov chains. Chapter two overviews the data and includes an initial analysis of the data set, describing the trend and nature of student population, graduation and dropouts at ETSU. In chapter three, we develop the initial transition matrices which describe the Markov chain model for students who transition from one semester to the next or who dropout or graduate. Chapter four provides a detailed explanation about the students who take one or more semesters off before returning to ETSU (as a state we call hiatus). We then incorporate this information into the final Markov chain model. In chapter five, we perform an analysis of incoming students which will be included in the simulation of our model discussed in chapter six. We conclude in chapter seven with a summary and future work.

2 DESCRIPTION AND PRELIMINARY ANALYSIS OF THE DATA SET

2.1 Data Description

This thesis uses data obtained from the office of institutional research ETSU on student enrollment from Fall 2008 to Spring 2017. The data set provides a multitude of information about the students. Each student had a randomly generated student ID, so the student's identity was not revealed. For each student, there were approximately 60 different factors tabulated, some of which are summarized in Table 1. Over 195,000 students are included in this data set. For the purpose of developing a discrete-time Markov chain model that describes the transition in student enrollment, only a handful of the 60 parameters were needed. We note that the excluded factors in the data set could possibly affect student enrollment and may be useful in a different type of predictive model; however, those same factors are not necessary in developing the type of Markov chain model considered in this thesis. We are only interested in students as they transition from freshman through senior including graduating or dropping out as well as those who take some time off from ETSU before returning. This implies graduate students or post graduates were left out as well as those who enter the college as undergraduate or graduate to pursue special programs which do not lead to a final degree. The parameters included in this model development include the student ID, the class level of the student, the semester or term of enrollment and whether or not the student has graduated during the semester.

Table 1: Description of Factors in the Data Set

Factors	Explanation and Description
Student ID	The ID number from each student
Academic period	The current academic year of student
Year term	Fall or Spring semester with the year
Year of registration	The year student registered into the program
Student level	Undergraduate or graduate
Reporting degree	Student degree upon admission
Graduated	Yes or No
Age during term	The age of student during corresponding term
Gender	Male or Female
Resident	In-state or out-of-state resident
Lottery resident	In-state or out-of-state resident defined by lottery scholarship
Zip code	Student resident zip code
Citizenship	Whether or not student is U.S. citizen
Previous registration	Pre-college student, first time student, returning student or readmitted student
Class Level	Freshman, Sophomore, Junior, or Senior
Registration type	The student registered as transferred student, transient, first time college, first time graduate or professional student and or student not classified in the categories given
Total term hour	The total credit hours student has earned in the semester
ACT score	The score student received in the ACT test
Ethnicity	White, Black, Hispanic, Asian, American, non-resident or unknown

2.2 Trend in Student Enrollment

We first looked at the trends exhibited in the data. To do this, the data was filtered and sorted into each term. Each semester was further sorted into class levels (freshman, sophomore, junior and senior) and then copy and pasted in Minitab. Figures 1-4 show the trends.

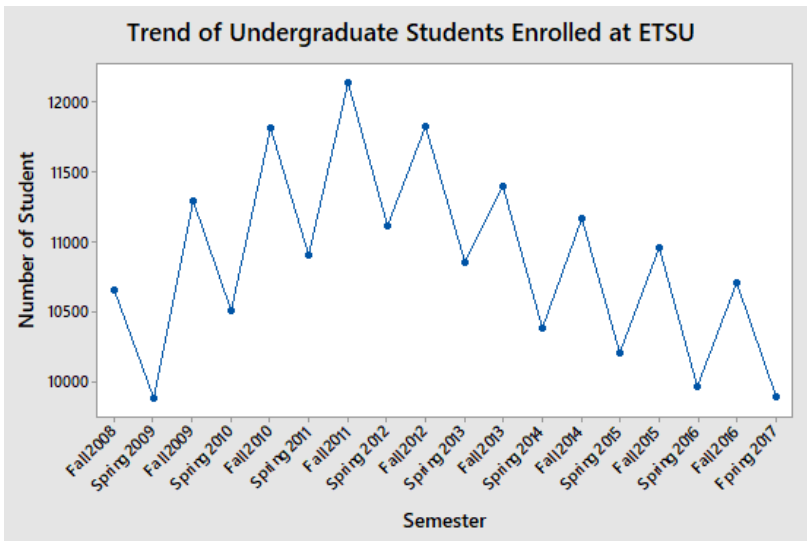


Figure 1: Trend of undergraduates student enrollment.

Figure 1 above shows the trend of undergraduate (freshman, sophomore, junior and senior) students at ETSU from fall 2008 to spring 2017 excluding summer semester. The output shows more students are enrolled in fall semesters than spring semesters across the course of the data. Indeed, this is most likely due to more students dropping out in spring semester than in fall semester as seen in Figure 4. There was an upward trend of students from 2008 to 2011 and thereafter enrollment continued to decline. “Historically, as the economy improves and Americans get back to work, college enrollment declines,” says U.S. Under Secretary of Edu-

cation Ted Mitchell [2]. For instance, considering the recent economic recession in the U.S. between December 2007 and June 2009, which saw many jobs shut down and laborers were laid off, resulted in an increased unemployment rate of 10% and a negative GDP of 5.1% [16]. Overall, enrollments of new students rose to slightly over 2 million in 2010, up from slightly less than 2 million in 2006, an increase of 6.8% [16]. Dewayne Matthews, vice president of strategy for the Indianapolis-based education nonprofit Lumina Foundation, outlined that College enrollment trends tend to be counter-cyclical in terms of employment, i.e. “When employment rates are low, enrollment goes up, and when employment rates are up, enrollment goes down.” [17] This mimics the trend in Figure 1 above.

We further break down the trend of total undergraduates by their class levels. One may ask which class level actually contributes most to the trend we are seeing in Figure 1. Figure 2 below depicts the trend of enrollment of undergraduate students by class. We can see from Figure 2 that the trend in freshman’s enrollment is most predominantly similar to that seen in Figure 1. Approximately 500-600 more freshman are enrolled during the fall semester than spring semester across the course of the data. The trend of sophomores and juniors enrolled are fairly uniform; however, there is a cyclic pattern, although less predominant than that seen in the freshman data, from fall to spring semester as well.

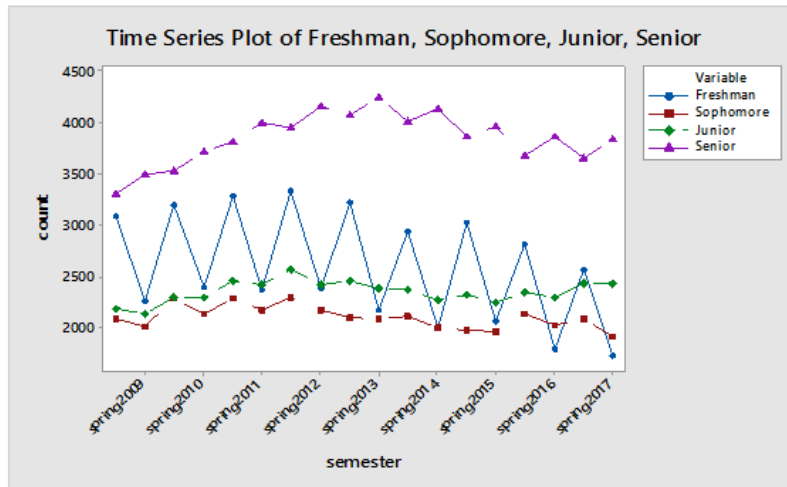


Figure 2: Trend of undergraduate student enrollment by class level.

However, the enrollment of students at the senior level in the Figure 2 shows a different trend from the other class levels. Here, more students are enrolled in the spring semester than the fall semester. This is most likely due to the fact that more students graduate during the spring semester thus less are enrolled the following fall semester. Figure 3 below shows the trend in the number of students graduating.

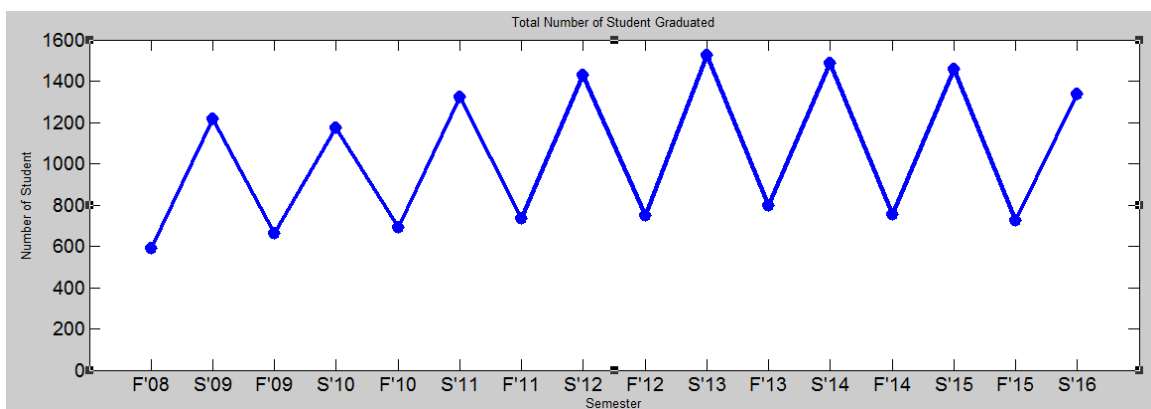


Figure 3: Trend of senior graduation.

It is also interesting to analyze students' dropout trends. Figure 4 which shows the trend in percentage of undergraduates who dropout by semester at ETSU. It can be seen that the percentage of students who dropout in spring semester is higher than for fall semester. The trend seen at ETSU is similar to that across the United States where thirty percent of college and university students drop out after their first year and as well as half never graduate, and college completion rates in the United States have been stalled for more than three decades [11]. Examining Figure 4, we found that in fall 2008, about 6% of students dropped out of ETSU. The percentage dropout is fairly stable across all fall semesters though the trend shows a slight growth in the dropout rate starting fall 2010 by about a 0.3% margin. The growth margin seem to continue throughout the rest of fall semester. There is a continuous increase in the dropout rate for spring semesters until spring 2012 where it attains the highest percentage dropout of about 13%. Most researchers on college student dropout in the U.S. have attributed the cause of the alarming rate of dropout as a result of economic growth and expansion [16]. The rate of dropout is troubling considering how well U.S. college educational systems are structured. However, "one cause of the dropout is also due to unprepared students signing up for school because they think a degree is their passport to the middle class" [13]. Just 56% of students who embark on a bachelor's degree program finish within six years, according to a 2011 Harvard study titled Pathways to Prosperity [13].

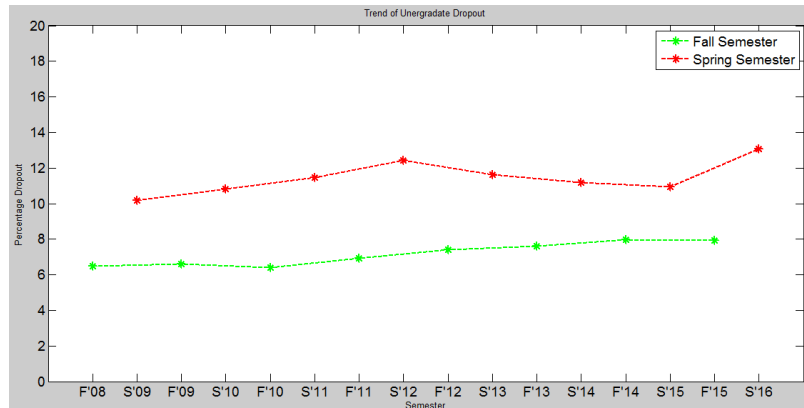


Figure 4: Trend of undergraduate student dropout.

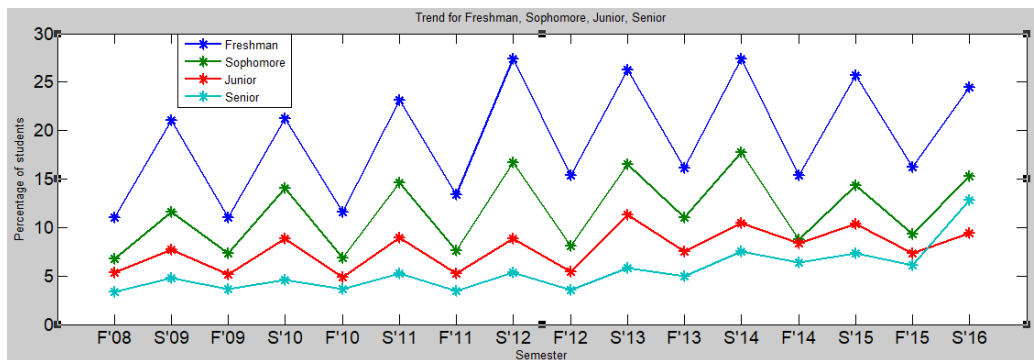


Figure 5: Trend of undergraduate student dropout for each class levels.

Figure 5 shows the trend for dropout percentages for freshmen, sophomores, juniors and seniors; In figure 5, we can see that freshmen represent the highest percentage of dropout in both fall and spring semesters, followed by sophomores, juniors and then seniors. In the past 20 years, more than 31 million students have enrolled in college, only to leave without a degree or certificate and a third of these students dropped out of school before the start of their sophomore year [12]. The first year on a college campus can be daunting for freshmen, with one out of three students not returning to that same institution for their sophomore year, according to report a

media group in Michigan [14]. This is why freshmen have the greatest dropout rate, and we see this same trend at ETSU.

3 MODEL DEVELOPMENT AND ANALYSIS

In this section, we discuss the development of the discrete-time Markov chain model using the data described in Section 2. As mentioned earlier, the Markov chain model is a stochastic process satisfying a memory-less property. We first analyze the model by visualizing a probability transition diagram for the model in Section 3.1 and then formalize the model in Section 3.2 with the probability transition matrices.

3.1 Discrete-Time Markov Chain Probability Transition Diagram

Figure 6 below depicts the transition state diagram for our Markov chain model for the progression of undergraduate students at ETSU. The transition probability diagram consists of seven (7) states, namely freshman, sophomore, junior, senior, dropout, graduation and hiatus. In the model, we assume there are only forward transitions and no backward transitions. For example, in the schematic in Figure 6, we assume, freshman can retain their freshman class level or progress to either a sophomore, junior, senior, graduate, drop out or take a break or hiatus from school. Sophomores either retains their status of sophomore or transition to junior, senior, graduate, drop out or hiatus, but sophomores cannot move back to freshman. Juniors and seniors follow a similar progression as freshmen and sophomores. However, some of the transitions are uncommon (such as freshman to senior), but they are included in the schematic because the data indicates these potential transitions (most likely due to erroneous classification of students within the data). Also, there are some backward transitions also indicated in the data which are believed to be errors in the incorrect categorization of transfer students initially. However, in the later years of

data collection, these backward transitions are infrequent. We disregard all backward transitions in the calculation of probability measurements. The hiatus state will be discussed in more detail in Section 4.

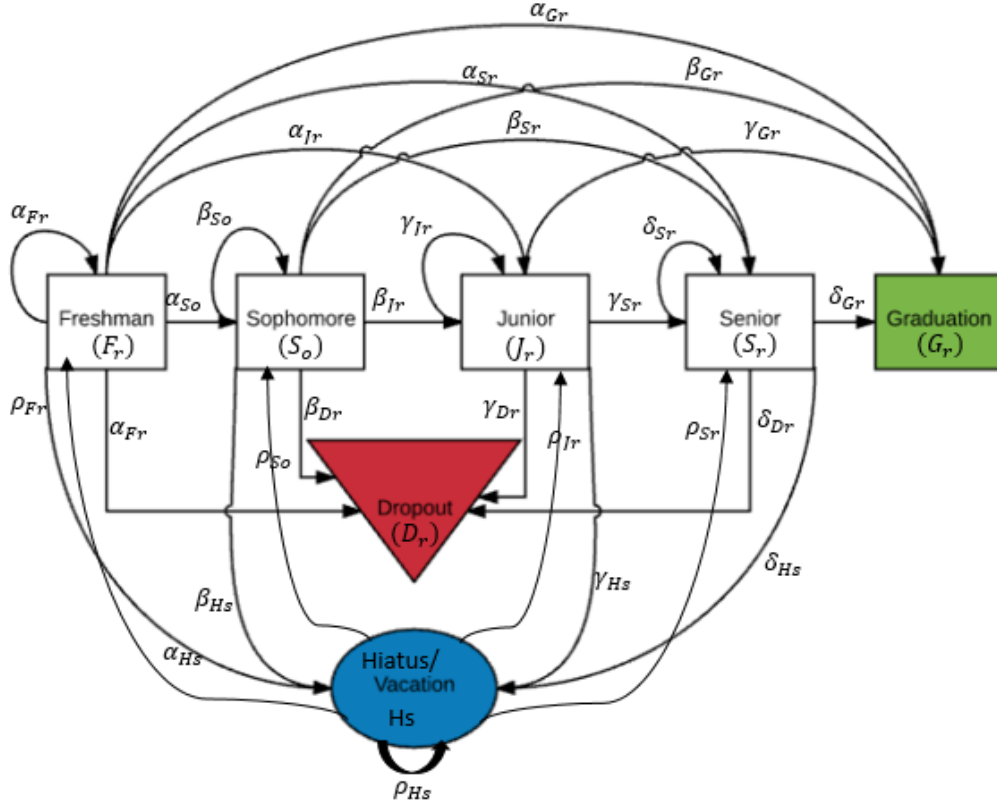


Figure 6: Discrete-Time Markov Chain Probability Transition Diagram

In the diagram α , β , γ , δ and ρ are the probabilities of transitions by freshmen, sophomores, juniors, seniors and hiatus state respectively to other states. Graduation and dropout states are said to be absorbing states. An absorbing state is a state that once entered, cannot be left; i.e. once a student enters this state, they are assumed to no longer transition to any other states within the system. The hiatus state is the only state that allows a student to enter, leave and re-enter or return. This can

be described as a ‘pendulum state’. It allows for free entry and exit of other states. However, we assume that there is no direct transition from the hiatus state to any of the two absorbing states.

3.2 Discrete-Time Markov Chain Transition Probability State Matrices

Now, we build a general probability transition matrix which is constructed of four smaller matrices, P1, P2, P3, and P4,

$$P = \begin{pmatrix} P1 & P2 \\ P3 & P4 \end{pmatrix}. \quad (2)$$

A discrete-time Markov chain is represented by a probability transition matrix P where $p_{ij}(n) = Prob\{X_{n+1} = j | X_n = i\}$ is the probability that a student is in state j in semester $n+1$ given that he was in state i during semester n , $i, j = 1, 2, \dots, N$, where N is the total number of semesters (i.e $N = 18$). The entry $p_{ij} \geq 0$ and $\sum p_{ij} = 1$. Considering matrix P in Equation (2), P1 represents the college to college probability transitions (e.g., freshmen to sophomore, junior, senior, dropout, and graduate but not to the hiatus state), P2 represents the transitions from the hiatus state to any of the college states (freshman, sophomore, junior, senior, dropout, and graduate), P3 the represents college to hiatus probability transitions and P4 represents college to hiatus probability transition matrix. Note that P1 and P3 are complements and the same for P2 and P4. Therefore, the sum of the first column of P1 and first column of P3 will equal 1. For example, all probabilities of the transition of freshmen in P1 plus probabilities of the transition of freshmen in P2 equal to 1. This is the same for P2 and P4. We discuss the formulation of P1 in this section and P2, P3 and P4 in

Section 4. The matrix P1 has the form.

$$\begin{array}{c}
 \text{Current - State} \\
 \begin{array}{cccccc}
 Fr & So & Jr & Sr & Dr & Gr \\
 \alpha_{Fr} & 0 & 0 & 0 & 0 & 0 \\
 \alpha_{So} & \beta_{So} & 0 & 0 & 0 & 0 \\
 \alpha_{Jr} & \beta_{Jr} & \gamma_{Jr} & 0 & 0 & 0 \\
 \alpha_{Sr} & \beta_{Sr} & \gamma_{Sr} & \delta_{Sr} & 0 & 0 \\
 \alpha_{Dr} & \beta_{Dr} & \gamma_{Dr} & \delta_{Dr} & 1 & 0 \\
 \alpha_{Gr} & \beta_{Gr} & \gamma_{Gr} & \delta_{Gr} & 0 & 1
 \end{array}
 \end{array}
 \begin{array}{l}
 Fr \\
 So \\
 Jr \\
 Sr \\
 Dr \\
 Gr
 \end{array}
 \text{Next - State} .$$

In the transition matrix P1, $p_{11} = \alpha_{Fr}$ is the probability that a freshman at the current time/semester remains a freshman at the next time step. $p_{12} = \alpha_{So}$ is the probability that a freshman during the current semester transitions to a sophomore during the next semester, etc. The sum of all probabilities involving transitioning by a freshman to other states is equal to one. This is the same for all transitions by other class states.

To calculate the probabilities in the transition matrix we start with the randomly generated student id in the data. We have 18 semesters worth of data. We initially considered all semesters separately which resulted in 17 probability transition matrices initially for analysis, and then the data is combined to form a single fall semester probability transition matrix and a single spring semester probability transition matrix.

The data set in Excel was filtered and sorted according to semester and then recoded for ease in use. We coded 1 for if the student has graduated and 0 if the student has not graduated. The class levels were coded 1 for freshman, 2 for sophomore, 3 for junior and 4 for senior. This Excel file was saved and imported in Matlab. In Matlab,

each student id was use to find the student in each class during a semester and track their class in the next semester. If a student is not found in the data for the next semester, we check whether the student has either graduated, taken a break (hiatus) or dropped out. We then divided the number of students in each state by the total number of students in the given class for the semester. For instance, α_{Fr} = number of freshmen next semester / total number of freshmen this semester, i.e. we obtain the total number of freshmen this semester who are still freshmen next semester and divide by the total number of freshmen this semester. The same calculation was made for all entries in the matrix.

The following are four of the 17 resulting probability transition matrices of P1; fall 2009 to spring 2010, spring 2010 to fall 2010, fall 2010 to spring 2011 and spring 2011 to fall 2011.

$$P1_{F'09S'10} = \begin{pmatrix} 0.633 & 0 & 0 & 0 & 0 & 0 \\ 0.220 & 0.550 & 0 & 0 & 0 & 0 \\ 0.004 & 0.337 & 0.567 & 0 & 0 & 0 \\ 0.003 & 0.007 & 0.352 & 0.746 & 0 & 0 \\ 0.110 & 0.073 & 0.052 & 0.036 & 1 & 0 \\ 0.000 & 0.001 & 0.003 & 0.190 & 0 & 1 \end{pmatrix} \quad (3)$$

$$P1_{S'10F'10} = \begin{pmatrix} 0.342 & 0 & 0 & 0 & 0 & 0 \\ 0.396 & 0.404 & 0 & 0 & 0 & 0 \\ 0.001 & 0.403 & 0.405 & 0 & 0 & 0 \\ 0.003 & 0.011 & 0.466 & 0.594 & 0 & 0 \\ 0.213 & 0.140 & 0.088 & 0.046 & 1 & 0 \\ 0.002 & 0.004 & 0.010 & 0.320 & 0 & 1 \end{pmatrix} \quad (4)$$

$$P1_{F'10S'11} = \begin{pmatrix} 0.630 & 0 & 0 & 0 & 0 & 0 \\ 0.229 & 0.540 & 0 & 0 & 0 & 0 \\ 0.002 & 0.356 & 0.568 & 0 & 0 & 0 \\ 0.001 & 0.004 & 0.360 & 0.754 & 0 & 0 \\ 0.116 & 0.068 & 0.049 & 0.036 & 1 & 0 \\ 0.001 & 0.001 & 0.001 & 0.182 & 0 & 1 \end{pmatrix} \quad (5)$$

$$P_{1_{S'11}F'11} = \begin{pmatrix} 0.348 & 0 & 0 & 0 & 0 & 0 \\ 0.382 & 0.410 & 0 & 0 & 0 & 0 \\ 0.000 & 0.408 & 0.415 & 0 & 0 & 0 \\ 0.001 & 0.004 & 0.456 & 0.579 & 0 & 0 \\ 0.232 & 0.146 & 0.090 & 0.053 & 1 & 0 \\ 0.001 & 0.001 & 0.009 & 0.334 & 0 & 1 \end{pmatrix} \quad (6)$$

It can be seen that there is not much variation for all the probabilities in the transition matrices for fall to spring semester; all spring to fall semester probability transition matrices are similar as well. For instance, in fall 2009 to spring 2010 (F'09S'10) the probability of transitioning from freshman to freshman is 63.3%, freshman to sophomore is 22%, freshman to junior is .4%, freshman to senior is .3%, freshman to dropout is 11% and 0% for freshman to graduation. These probabilities are virtually the same as that of fall 2010 to spring 2011. Spring 2010 to fall 2010 also has almost the same probabilities as spring 2011 to fall 2011. This runs through almost all of the 17 matrices with a few exceptional cases.

A student cannot graduate as a freshman, sophomore, junior or senior, but the data presents such instances in which this happened. We attribute this to errors in the data set. For example, in spring 2010 there is .3% and .2% of freshmen who transition to senior and graduation respectively. The matrices show further evidence that more students graduate or dropout in spring semester than fall semester. In spring 2011, 23.2% of freshmen dropout compared to 11.6% in fall 2010. Similarly, 33.4% of seniors graduated in spring 2011 as compare to 18.2% who graduated in fall 2011. This also corresponds to the data seen in figure 3 and 5 above, which showed the trends for graduation and dropouts respectively.

Now, after calculating and analyzing the 17 individual semesters which include

9 fall to spring semesters (2008 to 2016) and 8 spring to fall semesters for P1 (2009 to 2017) matrices, we then combined all fall semester data and all spring semester data to form one probability transition matrix for fall semester and one for spring semester. We used the same method as discussed previously; however, we did this on combined data for all fall and spring semesters. For instance, for the fall semester probability transition matrix, the probability of freshman to freshman is the addition of the number of all freshman to freshman for fall to spring semesters divided by the total number of freshmen for all fall to spring semesters. Below are the two resulting matrices for fall and spring semester;

$$P1_{AllFall} = \begin{pmatrix} 0.603 & 0 & 0 & 0 & 0 & 0 \\ 0.238 & 0.537 & 0 & 0 & 0 & 0 \\ 0.004 & 0.357 & 0.552 & 0 & 0 & 0 \\ 0.002 & 0.005 & 0.368 & 0.747 & 0 & 0 \\ 0.127 & 0.072 & 0.053 & 0.037 & 1 & 0 \\ 0 & 0.002 & 0.002 & 0.189 & 0 & 1 \end{pmatrix} \quad (7)$$

$$P1_{AllSpring} = \begin{pmatrix} 0.327 & 0 & 0 & 0 & 0 & 0 \\ 0.397 & 0.413 & 0 & 0 & 0 & 0 \\ 0.001 & 0.406 & 0.407 & 0 & 0 & 0 \\ 0.003 & 0.006 & 0.473 & 0.560 & 0 & 0 \\ 0.229 & 0.136 & 0.084 & 0.059 & 1 & 0 \\ 0.001 & 0.001 & 0.006 & 0.348 & 0 & 1 \end{pmatrix}. \quad (8)$$

The probability of transitioning between the various states doesn't change much compared to the individual fall and spring semester matrices with more students graduating or dropping out in the spring than the fall. Also, more freshmen dropout at the end of each semester as well as more seniors graduate in each semester. Examining the probability transition matrices in Equations (7) and (8), 12.7% and 22.9% of freshmen dropout during the fall semester and spring semester respectively. The percentage of seniors who graduated in fall and spring is 18.9% and 34.8% respectively.

The percentage of seniors who dropout within this period is 3.7% for fall and 5.9% for spring. Even though this percentage is not large, these percentages signify a decent number of students who have worked for several years to gain senior status and then not finish. In November 2014, it was reported that the U.S. has the lowest college completion rate in the developed world, at least among the 18 countries tracked by the Organization for Economic Cooperation and Development [13, 18]. Also, about 53% of undergraduates who attended full time or part time struggled with one-third dropping out entirely, others finished in six years and those who spend more than six years have a greater chance of dropout [18].

3.2.1 Matrix Comparison

In this section, we compare the probability transition matrices for fall and spring to see whether there is difference between them. We hypothesized that H_0 : there is no difference in the two matrices (null hypothesis) and H_1 : there a is difference in the two matrices (alternate hypothesis). We tested for the difference in the matrix using a pair sample t-test. The decision rule is that we reject H_0 if the $p - value$ is less than an α level of significance of 0.05 and conclude on H_1 .

First, we compare the various individual fall matrices to the probability transition matrix for fall using all the data. We tested for whether there is a difference between Fall 2009 and the all-data Fall matrix. This resulted to a $p - value = 0.8672$; hence we fail to reject H_0 and conclude there is no significant difference in the two matrices. We performed another test to see whether Fall 2010 matrix is different from Fall matrix using all the data. We had the same conclusion with a $p - value = 0.9099$. In

general, we conclude that the fall matrix using all the data is sufficient for modeling purposes.

Next, we performed the test for whether the individual spring matrices are different from the spring matrix using all the data. When we compare Spring 2010 and Spring 2011 to the Spring matrix using all the data in which we had p - values of 0.9034 and 0.6934 respectively. Hence, we also fail to reject the null hypothesis and conclude that the spring matrix using all the data is sufficient for modeling purposes.

Note that another way to check for whether the matrices are different is finding the eigenvalues of their difference. If the eigenvalues are positive it means the two matrices are not different. If the resulting eigenvalues of their difference are negative, we conclude the two matrices are different. We reached the same conclusion as using the t-test procedure.

We concluded statistically that there is no difference between the matrices using individual semester data for fall verses that using all the data for a fall matrix or individual spring verses the spring matrix using all the data or fall verses the spring semester matrices. However, although there is no significant difference, we can see the difference in Section 2 in the trends graphically when comparing spring and fall semesters and in the elements of the matrix given above. For this reason, it is appropriate to ‘measure the distance’ between the matrices. To do this, we find the matrix norm of the differences of the matrices.

A matrix norm of a matrix $\|P\|$ is any mapping from $R^{n \times n}$ to R with the properties:

- $\|P\| > 0$ if $P \neq 0$,

- $\|\alpha P\| = |\alpha| \|P\|$ for any $\alpha \in R$ and
- $\|P1 + P2\| \leq \|P1\| + \|P2\|$ (triangular inequality) for any matrix $P1, P2 \in R^{n \times n}$.

There are many types of matrix norms but for the purpose of this thesis we consider just two of them, i.e. the matrix 2-norm and the Frobenius norm. The induced **2-norm** is defined by

$$\|P\|_2 = s_1$$

where s_1 is largest singular value of P. On the other hand, the Frobenius norm of $P \in C^{m \times n}$ is defined as:

$$\|P\|_F^2 = \sum_{ij} |p_{ij}|^2 = \sum_i \|P_{i*}\|_2^2 = \sum_j \|P_{*j}\|_2^2 = \text{trace}(P \times P)$$

Note that $\|P\|_2 \leq \|P\|_F$

The above two types of matrix norms gives different results, but the conclusions are similar; therefore, we use the matrix 2-norm. For example, we compare Fall 2009 and Fall 2010 probability transition matrices with the fall semester matrix using all the data for fall semester by taking the matrix norm of their difference. The matrix norm of the difference for the Fall 2009 probability transition matrix and the fall semester matrix using all the data is 0.04 and between the Fall 2010 probability transition matrix and the fall semester matrix using all the data is .03. Similar calculations for Spring 2010 and Spring 2011 probability transition matrices verses the spring semester matrix using all the data resulted to a matrix norm of 0.04 and 0.037 respectively. The matrix norm of the difference between the fall and spring semester probability transition matrices resulted to 0.337. We can clearly see that fall semester matrices

are similar to one another since their matrix norms are close implying there is little difference among them. However, fall and spring semester matrices have a larger norm implying there is a larger difference between them. Therefore, for the purposes of modeling enrollment at ETSU, we choose to model spring and fall transitions separately.

4 DISCUSSION AND ANALYSIS OF THE HIATUS STATE

Hiatus literally means to take a break. In the course of the initial analysis of the data set, we discovered that there are students who are enrolled one semester, not enrolled the next semester but enrolled in a future semester (two or more semesters after the current semester of analysis). In other words, the student ‘takes a break’ from East Tennessee State University for one or more semesters. The class status of the student (freshman, sophomore, junior and senior) upon leaving may or may not change upon returning. Some students come back to the same class while others progress in their class status. We assume that some students may take a break for a semester or more to earn money to pay for college or for other reasons while others may transfer to another university, take some classes, and then return to ETSU in a later semester. If enough credits are earned, the student could change their class status. In the thesis by Helbert [10], such students were classified as having dropped out; it was assumed that once students left ETSU, they never returned [10]. This was a misclassification and could lead to the inflation of dropout. In order to correct the issue of misclassification, we include the hiatus state.

There are many reasons why a student may want to take a semester or more off from particular university. Burnout is one of the top reasons students take a break, i.e. a student may take break to regain focus after a long term of working too hard. At ETSU, it is also hypothesized that students sometimes take time off to earn money to support themselves. It is advisable for a student faced with problems of finance such as the inability to pay tuition and ordinary expenses to take a break rather than accumulating student loans which may jeopardize their career success. This

way the student can work for some time to take care of the expenses accrued. As a student, one should also be able to validate his/her career option. Having a doubt about one's career choice along the way may require taking a break. This can give the student some valuable hands-on experience in the chosen field and help develop more a realistic idea of what the career field is like, and help him/her validate his/her goals. Having physical and emotional health problems can make your college life miserable and prevent you from doing your best in school. If the problem is starting to get in the way of your studies, taking a break and giving oneself some time to heal may be the best option [15].

Any student who returns back to ETSU after leaving a particular semester is classified in the hiatus state in our model. In our model, hiatus is the only state that allows for transitions both in and out of the state. Thus, a student can enter hiatus more than once. The description of the hiatus state has been summarized in Table 2 below.

Table 2: A table showing the description of variables to analyze for students taking a break

Explanation / Description
Students ID
The semester the student left on hiatus
The semester the student returns after hiatus
The number of semesters the student spent out of ETSU on hiatus
The class level of the student before hiatus
The class level of the student after hiatus
The number of class levels the student has progressed while on a break

There are eighteen semesters in total from fall 2008 to spring 2017. A student can leave from the first (fall 2008) through the sixteenth semester (fall 2016) and return from the third (fall 2009) through the eighteenth semester (spring 2017). Our analysis does not include or account for students who left fall 2016 since we cannot anticipate when they will return. This enables us to obtain the number of semesters the student spends in hiatus before returning to ETSU. For example, if a student is enrolled in fall 2008 and then not enrolled again until fall 2009, the student has spent only one semester on hiatus. To obtain the class level upon returning from hiatus, we track the status of the student before and after hiatus. For example, if the student was a freshman before hiatus and still returns as a freshman after hiatus, the student made no transition and still remains in the same class. If the student was a sophomore and returned as a junior, then the student has made a transition by one class level. Hence, zero represents students who maintain their class levels or status after hiatus, one represents students who made one transition to the next level, say from freshman to sophomore or from junior to senior and two represents students who made a double move, say from freshman to junior or from sophomore to senior. The jump in class status after a student has taken a break from ETSU is possible as a result of transfer credit hours. We do not expect a student to make a backward transition after hiatus; in other words, a student cannot return at a class level lower than that which they left (e.g. return as a freshman when they left as a sophomore).

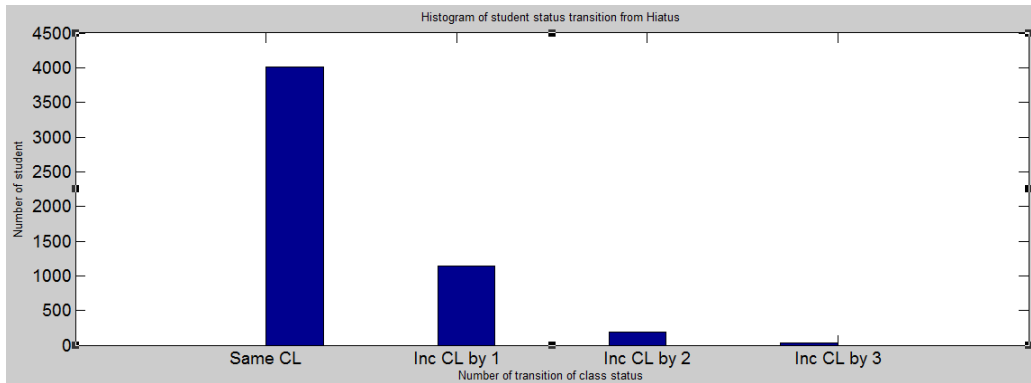


Figure 7: Number of students who transitions to the various classes after hiatus

There are 5,394 students total who took a break from ETSU for one or more semesters (i.e. transition to a hiatus state within the model), between fall 2008 and fall 2016 (within the 16 semesters). Out of this number, 1374 are freshmen representing about 25.5%, 1118 are sophomores representing about 20.7%, 1044 are juniors representing about 19.4% and 1858 are seniors representing about 34.4%. In Figure 7, out of the number of students who took a break, 4,008 (74.3%) return to the same class status or level they left, 1,143 (21.2%) of them made a one step progression of class status, about 192 (4.6%) returned with a class status two levels above that which they left and 31 (.6%) made a three step transition (i.e freshman to senior). This is not unexpected based on the many reasons or circumstances for which students take a break as discussed above. Some students take a break due to burnout or from financial constraint; therefore, these students would not be expected to take classes at another institution and thus they return at the same class level. In addition, it is very rare or unusual to have a student who left the college in a particular semester as a freshman to return after some period as a senior. Therefore,

it is not surprising this transition is the lowest.

When a student is on hiatus, how many semesters do they actually spend away from ETSU? Figure 8 gives a boxplot showing all sixteen semesters in which students left at ETSU and the number of semesters they spent away before returning. We can infer from the plot that on average a student who leaves after fall semester spends a median of only one semester off before returning. For example, a student who left after fall 2008 might not enroll for spring 2009 but returns in fall 2009. In contrast, the median number for students who take a break after the spring semester is two i.e. a student who leaves after spring semester (who doesn't return the following fall semester), will most likely take an entire year off before returning to ETSU. For instance, a student who takes a break after spring 2009 is most likely to return in fall 2010, thereby spending fall 2009 and spring 2010 on hiatus. The boxplots show a decreasing number of outliers in the data for this analysis. This is most likely due to the fact, that there is a longer period for which the student might return in the earliest data.

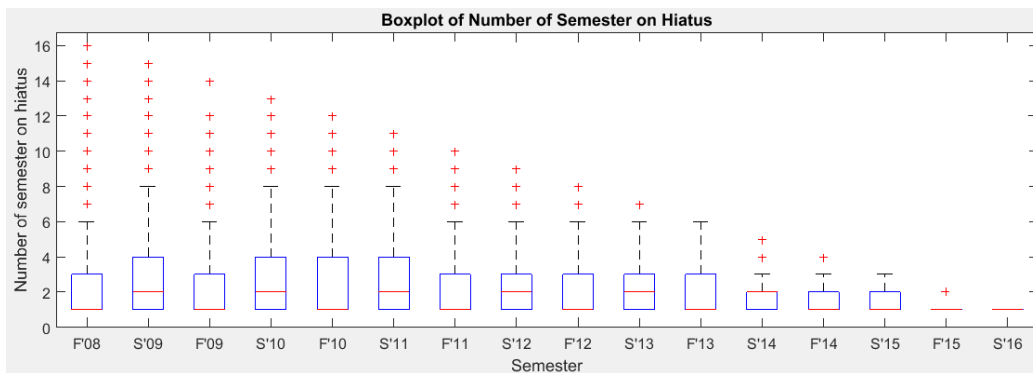


Figure 8: Side by side boxplot on the number of semester spent on hiatus

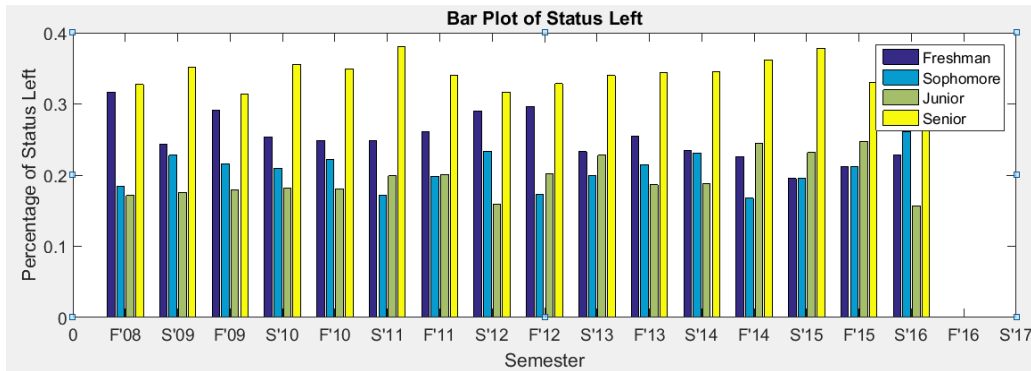


Figure 9: Bar plot of student status left

Figure 9 shows the percentages of students who took break across the various class levels in a particular semester. For instance, in spring 2016 about 20% of student who left were freshmen, about 25% were sophomore, about 15% were junior and about 35% were senior. Seniors constitute the majority who took the break, followed by freshman, sophomore and then junior.

It is also necessary to consider the class status/level of a student when returning from hiatus. This is because not all students' maintain their status whilst returning from break. Figure 10 shows students class status after taking a break. For example, of the students who took a break in Fall 2008, 42% return as a senior, 20% return as a junior level, 20% return as sophomore and 18% returned as freshmen.

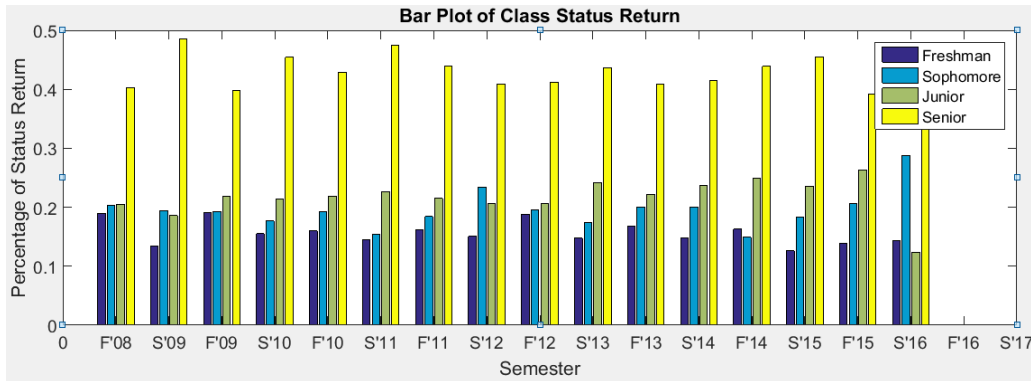


Figure 10: Bar plot of student status Return

We are probably more interested in knowing to which class each class of students return. When we consider freshmen, we may be interested in knowing what percentage of them return as freshman or transition to other class level. We also do the same for sophomores and juniors as well. Figures 11-13 show the class status of students in ETSU after taking a break for freshman, sophomore and junior respectively for each of the 18 semesters data. In Figure 11, out of the Freshmen who returned from hiatus in Fall 2008, 60% returned as Freshmen, 23% returned as Sophomores, 15% returned as Junior and 2% returned as Senior. It can be seen that about an average of 62% of Freshmen who take break for one or more semesters in ETSU return as freshmen, an average of 25%, 10% and 3% returned as sophomores, juniors and seniors respectively. In Figure 12, in Fall 2008, 70% of sophomores who took a break returned as Sophomore, 22% returned as junior and about 8% returned as Seniors. On average 60% of sophomores do return as sophomore after taking a break. Figure 13 shows that most juniors who take a break return as junior with an average of 70%. Spring 2009 was an exceptional situation in which the same number of students returned to junior

as the number of those returning as seniors. In general, we expected more students to return to the same class they left, and the diagrams clearly show that. We know that a senior who takes a break must return as a seniors. There are no backward transitions and no student can enter any of the two absorbing states (graduation and dropout) from hiatus.

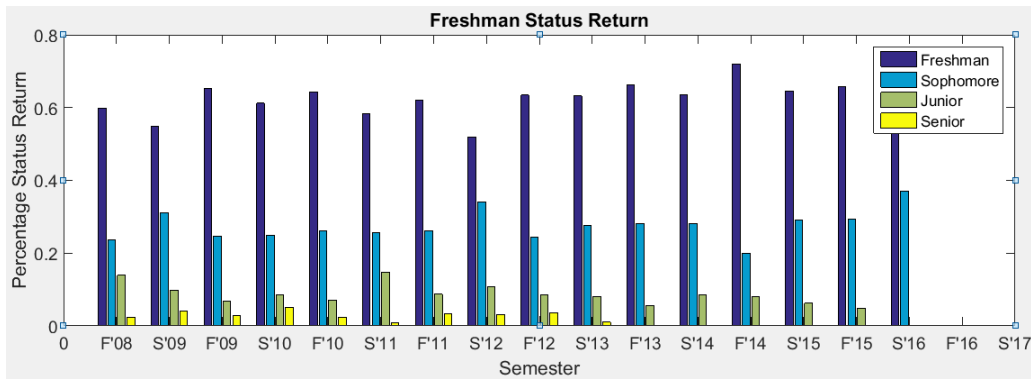


Figure 11: Bar plot of Freshman status Return

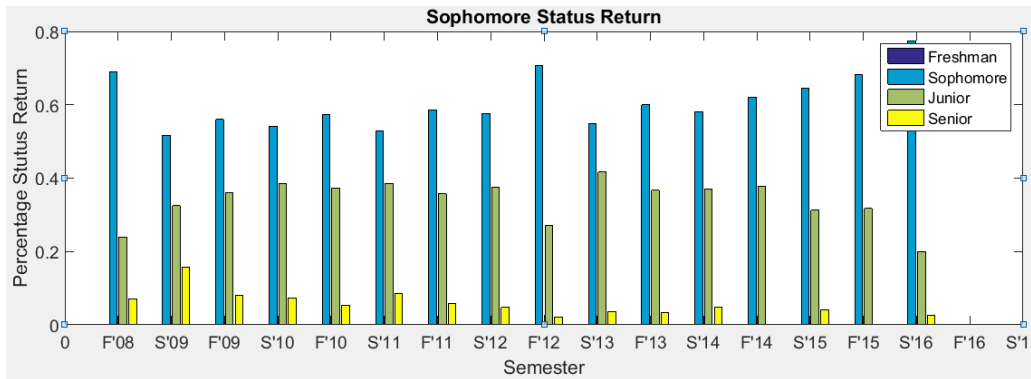


Figure 12: Bar plot of Sophomore status Return

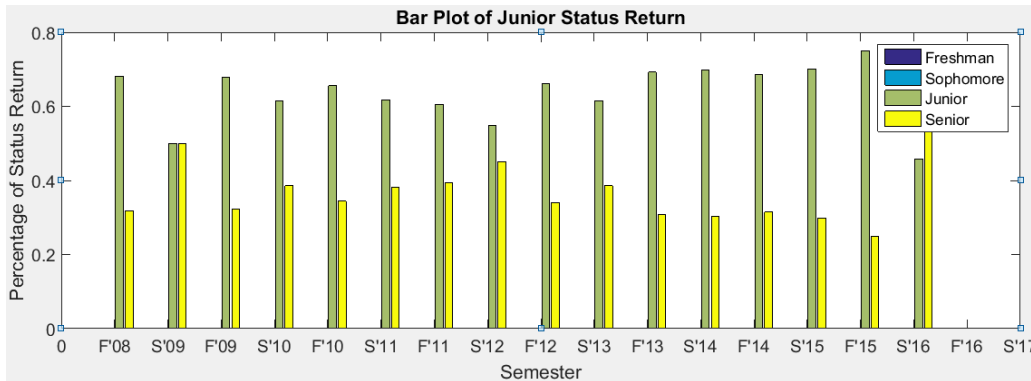


Figure 13: Bar plot of Junior status Return

4.1 Development of the Hiatus Model

In this section, we expand the hiatus transition state in Figure 6 to obtain a more precise probability transition diagram for the hiatus state. We also consider the other transition probability matrices P2, P3 and P4 in the general probability transition matrix P in Equation (2). These matrices show which class level a student transitions to after taking a break (P2) or probability transitions from ETSU to hiatus (P3) or from hiatus to hiatus (P4). The three matrices sum up our analysis on the hiatus state. Figure 14 below represents the probability transition diagram for the hiatus state where α , β , γ and δ represent the transition probabilities of students for freshman, sophomore, junior and senior respectively to other states and ρ represents probability of transition from the hiatus state to other states. ρ_{HF} , ρ_{HSo} , ρ_{HJr} and ρ_{HSr} represent freshmen, sophomores, juniors and seniors who are on break.

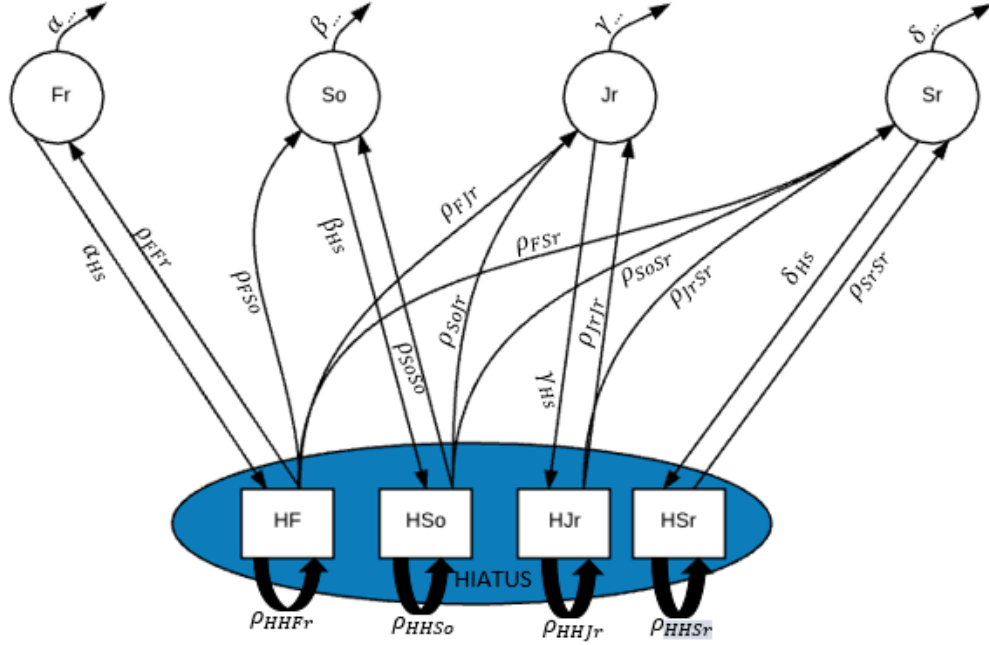


Figure 14: The transition probability diagram for hiatus.

We consider matrices P2, P3 and P4 for fall and spring semesters using all the data and then merge them with P1 in Section 3 to formulate our general matrix P in Equation (2). The matrix P2 has the form

$$P2 = \begin{matrix} & \begin{matrix} HFr & HSo & HJr & HSr \end{matrix} \\ \begin{matrix} HFr \\ HSo \\ HJr \\ HSr \\ Dr \\ Gr \end{matrix} & \begin{pmatrix} \rho_{FrFr} & 0 & 0 & 0 \\ \rho_{FrSo} & \rho_{SoSo} & 0 & 0 \\ \rho_{FrJr} & \rho_{SoJr} & \rho_{JrJr} & 0 \\ \rho_{FrSr} & \rho_{SoSr} & \rho_{JrSr} & \rho_{SrSr} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix} \begin{matrix} Fr \\ So \\ Jr \\ Sr \\ Dr \\ Gr \end{matrix} \text{ Next - State , }$$

where ρ_{FrFr} represents the probability that a freshman who took a break from ETSU as a freshman (HF) will return as a freshman. Similarly, ρ_{FrSo} represents the

probability that a student who took a break as a freshman will return as a sophomore, etc. It is assumed that no student transitions directly from hiatus to either dropping out or graduating; therefore, these entries are all 0. Also, there are no backward transitions; therefore, the matrix is upper triangular. Equations (9) and (10) are the resulting matrices for P2 for fall and spring semesters, respectively. To calculate the entry ρ_{FrFr} we divide the number of freshmen who took a break and returned as a freshmen by the total number of students who took a break as a freshman. Similarly, ρ_{FrSo} is equal to the number of freshmen who took a break and returned as a sophomore divided by the total number of students who took a break as a freshman state.

$$P2_{AllFall} = \begin{pmatrix} 0.298 & 0 & 0 & 0 \\ 0.114 & 0.298 & 0 & 0 \\ 0.001 & 0.160 & 0.305 & 0 \\ 0 & 0.006 & 0.175 & 0.478 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (9)$$

$$P2_{AllSpring} = \begin{pmatrix} 0.406 & 0 & 0 & 0 \\ 0.142 & 0.431 & 0 & 0 \\ 0.006 & 0.219 & 0.407 & 0 \\ 0.003 & 0.004 & 0.212 & 0.552 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (10)$$

In P2, the probability that a freshman on break in a fall semester returns as a freshman is 29.8%, the probability that a freshman returns as a sophomore is 11.4%; there is 0.1% probability that a freshman returns as a junior and 0% probability that a freshman returns as a senior after hiatus. Also, the probability that a freshman on break in the spring semester returns as a freshman is 40.6%, return as a sophomore 14.2%, 0.6% as a junior and 0.3% as a senior. It can be seen that there is a higher probability that a freshman after taking a break returns as a freshman in spring

semester than in fall semester. A similar analysis applies to sophomores, juniors and seniors.

The probability transition matrix of students transitioning from college (from the various class states) to the hiatus state is given by

$$\begin{array}{c}
 \text{Current - State} \\
 \begin{array}{cccccc}
 Fr & So & Jr & Sr & Dr & Gr
 \end{array} \\
 P3 = \begin{pmatrix}
 \alpha_{Hs} & 0 & 0 & 0 & 0 & 0 \\
 0 & \beta_{Hs} & 0 & 0 & 0 & 0 \\
 0 & 0 & \gamma_{Hs} & 0 & 0 & 0 \\
 0 & 0 & 0 & \delta_{Hs} & 0 & 0
 \end{pmatrix} \begin{array}{l}
 HFr \\
 HSo \\
 HJr \\
 HSr
 \end{array} \text{ Next - State}
 \end{array}$$

In the matrix P3, *Fr*, *So*, *Jr*, *Sr*, *Dr* and *Gr* represents the current states and the next state (the hiatus state) is represented by *HFr*, *HSo*, *HJr* and *HSr*. Also, α , β , γ and δ play a similar role as described in P2. Equations (11) and (12) are the resulting matrices for fall and spring semesters for P3. To calculate α_{Hs} , we divide the number of freshmen who took a break by the total number of freshmen; β_{Hs} is equal to the number of sophomores who took a break divided by the total number of sophomores. The same calculation was done for γ_{Hs} and δ_{Hs} . In fall semester (Equation (11)), there is 2.6%, 2.8%, 2.5% and 2.7% probability of a freshmen, sophomore, junior and senior taking a break respectively. For spring semester (Equation (12)), there is a 4.2%, 3.9%, 3.1% and 2.3% probability of freshmen, sophomores, juniors and seniors taking a break respectively. It can be seen that more students at ETSU take a break after spring semesters than fall semesters for freshmen, sophomores and juniors except seniors. Fewer seniors take break in spring semester than fall semester, because more seniors graduate in spring than fall as seen in Figure 3.

$$P3_{AllFall} = \begin{pmatrix} 0.026 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.028 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.025 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.027 & 0 & 0 \end{pmatrix} \quad (11)$$

$$P3_{AllSpring} = \begin{pmatrix} 0.042 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.039 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.031 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.023 & 0 & 0 \end{pmatrix} \quad (12)$$

Matrix P4 is the matrix representing the transition of students from hiatus to hiatus state, that is the probability of students who take a break and remain on the break the next semester. In this matrix, the state represents the class level of the student upon leaving ETSU; therefore, they will not transition to any other hiatus state than that which they are currently. The matrix P4 is given by

$$P4 = \begin{pmatrix} HF_r & HSo & HJ_r & HS_r \\ \rho_{HHF_r} & 0 & 0 & 0 \\ 0 & \rho_{HHSo} & 0 & 0 \\ 0 & 0 & \rho_{HHJ_r} & 0 \\ 0 & 0 & 0 & \rho_{HHS_r} \end{pmatrix} \begin{matrix} HF_r \\ HSo \\ HJ_r \\ HS_r \end{matrix} \quad \begin{matrix} Current - State \\ \\ \\ \\ \\ Next - State \end{matrix}$$

we calculate ρ_{HHF_r} by dividing the number of freshmen who took a break and remains on break the next semester by the total number of freshmen in hiatus state. Similar calculations are done for ρ_{HHSo} , ρ_{HHJ_r} and ρ_{HHS_r} . ρ represents probability of transition from Hiatus state to other states. The resulting matrix for fall and spring semester semesters are given in Equations (13) and (14) respectively. In fall semester, there is a 58.7%, 53.6%, 52.0% and 52.2% probability of freshmen, sophomores, juniors and seniors who took a break respectively still remains on break. For spring semester, there is a 44.3%, 34.4%, 37.1% and 44.7% probability of freshmen, sophomores, juniors

and seniors respectively still remains on break.

$$P4_{AllFall} = \begin{pmatrix} 0.587 & 0 & 0 & 0 \\ 0 & 0.536 & 0 & 0 \\ 0 & 0 & 0.520 & 0 \\ 0 & 0 & 0 & 0.522 \end{pmatrix} \quad (13)$$

$$P4_{AllSpring} = \begin{pmatrix} 0.443 & 0 & 0 & 0 \\ 0 & 0.344 & 0 & 0 \\ 0 & 0 & 0.371 & 0 \\ 0 & 0 & 0 & 0.447 \end{pmatrix} \quad (14)$$

By putting together P1, P2, P3 and P4 we get the general transition matrix (P) of the Markov chain model

$$P = \begin{pmatrix} Fr & So & Jr & Sr & Dr & Gr & HFr & HSo & HJr & HSr \\ \alpha_{Fr} & 0 & 0 & 0 & 0 & 0 & \rho_{FrFr} & 0 & 0 & 0 \\ \alpha_{So} & \beta_{So} & 0 & 0 & 0 & 0 & \rho_{FrSo} & \rho_{SoSo} & 0 & 0 \\ \alpha_{Jr} & \beta_{Jr} & \gamma_{Jr} & 0 & 0 & 0 & \rho_{FrJr} & \rho_{SoJr} & \rho_{JrJr} & 0 \\ \alpha_{Sr} & \beta_{Sr} & \gamma_{Sr} & \delta_{Sr} & 0 & 0 & \rho_{FrSr} & \rho_{SoSr} & \rho_{JrSr} & \rho_{SrSr} \\ \alpha_{Dr} & \beta_{Dr} & \gamma_{Dr} & \delta_{Dr} & 1 & 0 & 0 & 0 & 0 & 0 \\ \alpha_{Gr} & \beta_{Gr} & \gamma_{Gr} & \delta_{Gr} & 0 & 1 & 0 & 0 & 0 & 0 \\ \alpha_{Hs} & 0 & 0 & 0 & 0 & 0 & \rho_{HHFr} & 0 & 0 & 0 \\ 0 & \beta_{Hs} & 0 & 0 & 0 & 0 & 0 & \rho_{HHSo} & 0 & 0 \\ 0 & 0 & \gamma_{Hs} & 0 & 0 & 0 & 0 & 0 & \rho_{HHJr} & 0 \\ 0 & 0 & 0 & \delta_{Hs} & 0 & 0 & 0 & 0 & 0 & \rho_{HHSr} \end{pmatrix} \begin{matrix} Fr \\ So \\ Jr \\ Sr \\ Dr \\ Gr \\ HFr \\ HSo \\ HJr \\ HSr \end{matrix}$$

$$P_{AllFall} = \begin{pmatrix} Fr & So & Jr & Sr & Dr & Gr & HFr & HSo & HJr & HSr \\ 0.603 & 0 & 0 & 0 & 0 & 0 & 0.298 & 0 & 0 & 0 \\ 0.238 & 0.537 & 0 & 0 & 0 & 0 & 0.114 & 0.298 & 0 & 0 \\ 0.004 & 0.357 & 0.552 & 0 & 0 & 0 & 0.001 & 0.160 & 0.305 & 0 \\ 0.002 & 0.005 & 0.368 & 0.747 & 0 & 0 & 0 & 0.006 & 0.175 & 0.478 \\ 0.127 & 0.072 & 0.053 & 0.037 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.002 & 0.002 & 0.189 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0.026 & 0 & 0 & 0 & 0 & 0 & 0.587 & 0 & 0 & 0 \\ 0 & 0.028 & 0 & 0 & 0 & 0 & 0 & 0.536 & 0 & 0 \\ 0 & 0 & 0.025 & 0 & 0 & 0 & 0 & 0 & 0.520 & 0 \\ 0 & 0 & 0 & 0.027 & 0 & 0 & 0 & 0 & 0 & 0.522 \end{pmatrix} \begin{matrix} Fr \\ So \\ Jr \\ Sr \\ Dr \\ Gr \\ HFr \\ HSo \\ HJr \\ HSr \end{matrix}$$

for fall semester and for spring by

$$P_{AllSpring} = \begin{pmatrix} Fr & So & Jr & Sr & Dr & Gr & HFr & HSo & HJr & HSr \\ 0.327 & 0 & 0 & 0 & 0 & 0 & 0.406 & 0 & 0 & 0 \\ 0.397 & 0.413 & 0 & 0 & 0 & 0 & 0.142 & 0.431 & 0 & 0 \\ 0.001 & 0.406 & 0.407 & 0 & 0 & 0 & 0.006 & 0.219 & 0.407 & 0 \\ 0.003 & 0.006 & 0.473 & 0.560 & 0 & 0 & 0.003 & 0.004 & 0.212 & 0.552 \\ 0.229 & 0.136 & 0.084 & 0.059 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0.001 & 0.001 & 0.006 & 0.348 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0.042 & 0 & 0 & 0 & 0 & 0 & 0.443 & 0 & 0 & 0 \\ 0 & 0.039 & 0 & 0 & 0 & 0 & 0 & 0.344 & 0 & 0 \\ 0 & 0 & 0.031 & 0 & 0 & 0 & 0 & 0 & 0.371 & 0 \\ 0 & 0 & 0 & 0.033 & 0 & 0 & 0 & 0 & 0 & 0.448 \end{pmatrix} \begin{matrix} Fr \\ So \\ Jr \\ Sr \\ Dr \\ Gr \\ HFr \\ HSo \\ HJr \\ HSr \end{matrix}$$

Considering the matrix P for transitioning from fall semester to spring semester,, the probability that a freshman maintains the same class as freshman the following semester is 60.3%, probability that a freshman transitions to a sophomore is 23.8%; there is 0.4% probability that a freshman progresses to junior and 0.2% of a freshman progressing to the senior class. Also, the probability that a freshman dropouts is 12.7%; there is 0% probability of freshmen graduating and 2.6% probability that a freshman takes a break from fall to spring semester. The remaining states follow the same analysis as the freshman state.

4.2 Analysis and Long-Term Future Prediction

In this section, we look at where students transition after one year and five years. We also make long-term distribution of students to the various states. Now, we combine the spring and fall matrix of P to form a full one year probability transition matrix. This is calculated by multiplying the two matrices in order, i.e. spring P

matrix times fall P matrix. The one year matrix is of the form

$$P_{OneYear} = \begin{pmatrix} Fr & So & Jr & Sr & Dr & Gr & HFr & HSo & HJr & HSr \\ 0.208 & 0 & 0 & 0 & 0 & 0 & 0.336 & 0 & 0 & 0 \\ 0.341 & 0.234 & 0 & 0 & 0 & 0 & 0.249 & 0.355 & 0 & 0 \\ 0.099 & 0.369 & 0.235 & 0 & 0 & 0 & 0.051 & 0.303 & 0.341 & 0 \\ 0.006 & 0.175 & 0.472 & 0.433 & 0 & 0 & 0.004 & 0.083 & 0.353 & 0.556 \\ 0.298 & 0.175 & 0.121 & 0.082 & 1 & 0 & 0.084 & 0.054 & 0.036 & 0.028 \\ 0.002 & 0.006 & 0.133 & 0.449 & 0 & 1 & 0 & 0.003 & 0.063 & 0.166 \\ 0.037 & 0 & 0 & 0 & 0 & 0 & 0.273 & 0 & 0 & 0 \\ 0.009 & 0.030 & 0 & 0 & 0 & 0 & 0.004 & 0.196 & 0 & 0 \\ 0 & 0.011 & 0.026 & 0 & 0 & 0 & 0 & 0.005 & 0.202 & 0 \\ 0 & 0 & 0.012 & 0.037 & 0 & 0 & 0 & 0 & 0.006 & 0.250 \end{pmatrix} \begin{matrix} Fr \\ So \\ Jr \\ Sr \\ Dr \\ Gr \\ HFr \\ HSo \\ HJr \\ HSr \end{matrix}$$

In analyzing the one year probability transitions of students in ETSU, it can be seen that if a student starts as a freshman, there is 20.8% probability that the student remains a freshman, 34.1% probability that a freshman progresses to a sophomore, 9.9% probability that a freshman transitions to a junior and 0.6% probability that a freshman transitions to a senior after one year. A freshman has a 29.8% and 0.2% of dropping out and graduating respectively. The probability that a freshman progresses to a senior or graduates after a year is unexpected but these transitions are seen in the data most likely due to classification errors in the data. The reasons have been outlined on Section 3. It is important to recognize that a freshman has 3.7% probability of taking a break after 1 year and 0.9% probability of taking break as sophomore. Also, it can be seen that a freshman on hiatus has 33.6% probability to return as a freshman, 24.9% probability to return as a sophomore, 5.1% probability return as a junior, 0.4% of returning as a senior and 8.4% of dropping out. A freshman on hiatus has a 27.3% probability of remaining on break after 1 year. It is necessary to know that after one year, a student at the senior level has 43.3% of remaining a senior, 8.2% probability of dropping out, 44.9% probability of graduating and 3.7% probability of

taking a break. Next, we look at the behavior of the probability transitions of the states after 5 year below.

$$P_{5\text{years}} = \begin{pmatrix}
 & Fr & So & Jr & Sr & Dr & Gr & HFr & HSo & HJr & HSr \\
 Fr & 0.002 & 0 & 0 & 0 & 0 & 0 & 0.008 & 0 & 0 & 0 \\
 So & 0.010 & 0.002 & 0 & 0 & 0 & 0 & 0.028 & 0.006 & 0 & 0 \\
 Jr & 0.027 & 0.011 & 0.002 & 0 & 0 & 0 & 0.055 & 0.024 & 0.005 & 0 \\
 Sr & 0.102 & 0.092 & 0.055 & 0.027 & 0 & 0 & 0.139 & 0.133 & 0.086 & 0.057 \\
 Dr & 0.585 & 0.388 & 0.253 & 0.149 & 1 & 0 & 0.502 & 0.332 & 0.210 & 0.142 \\
 Gr & 0.260 & 0.495 & 0.683 & 0.820 & 0 & 1 & 0.246 & 0.489 & 0.687 & 0.793 \\
 HFr & 0.001 & 0 & 0 & 0 & 0 & 0 & 0.004 & 0 & 0 & 0 \\
 HSo & 0.002 & 0.001 & 0 & 0 & 0 & 0 & 0.004 & 0.001 & 0 & 0 \\
 HJr & 0.003 & 0.002 & 0 & 0 & 0 & 0 & 0.005 & 0.003 & 0.001 & 0 \\
 HSr & 0.008 & 0.010 & 0.007 & 0.004 & 0 & 0 & 0.010 & 0.012 & 0.010 & 0.008
 \end{pmatrix}$$

Most colleges expect or hope that most freshmen will graduate within five years. It is interesting to know that given the model, it shows that after 5 years, a freshman at ETSU has a 26.0% probability of graduating and 58.5% probability of dropping out. A sophomore has 38.8% probability of dropping out and 49.5% of graduating. If the student is a junior, then after five years the expected probability of dropping out is 25.3% and graduating is 68.3%. A senior level student may graduate at a probability of 82.0% or dropout at a probability of 14.9% after 5 years. Furthermore, we consider the long-term probability transition trend of the states below.

$$P_{50Years} = \begin{pmatrix} Fr & So & Jr & Sr & Dr & Gr & HFr & HSo & HJr & HSr \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.617 & 0.408 & 0.263 & 0.154 & 1 & 0 & 0.561 & 0.364 & 0.226 & 0.152 \\ 0.384 & 0.592 & 0.737 & 0.846 & 0 & 1 & 0.439 & 0.636 & 0.774 & 0.848 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} Fr \\ So \\ Jr \\ Sr \\ Dr \\ Gr \\ HFr \\ HSo \\ HJr \\ HSr \end{matrix} .$$

By removing the zeros, we are left with

$$P_{50Years} = \begin{pmatrix} Fr & So & Jr & Sr & Dr & Gr & HFr & HSo & HJr & HSr \\ 0.617 & 0.408 & 0.263 & 0.154 & 1 & 0 & 0.561 & 0.364 & 0.226 & 0.152 \\ 0.384 & 0.592 & 0.737 & 0.846 & 0 & 1 & 0.439 & 0.636 & 0.774 & 0.848 \end{pmatrix} \begin{matrix} Dr \\ Gr \end{matrix} .$$

In the very long-run, all other states transition to the two absorbing states (i.e. dropout and graduation) as shown in the 50 years probability transition matrix. This is also summarized in Table 3 below.

Table 3: Prediction of Long-Term Probability Transitions of the States

States	% Dropout (Dr)	% Graduation (Gr)
Freshman (Fr)	61.7	38.4
Hiatus Freshmen (HFr)	56.1	43.9
Sophomore (So)	40.8	59.2
Hiatus Sophomore (HSo)	36.4	63.6
Junior (Jr)	26.3	73.7
Hiatus Junior (HJr)	22.6	77.4
Senior (Sr)	15.4	84.6
Hiatus Senior (HSr)	15.2	84.8

According to the model, in the long-term, there is a 61.7%, 40.8%, 26.3% and 15.4% of a freshman, a sophomore, a junior and a senior at ETSU dropping out respectively. On the other hand, there is a 38.4%, 59.2%, 73.3% and 84.6% of a freshman, a sophomore, a junior and a senior graduating respectively. It is important to know that, according to the model, a student who takes a break at some point from any of the states has a lesser probability of dropping out and a higher probability of graduating as compared to a student who never had a break. For instance, a freshman who took a break has 5.5% more chance of graduating than the probability of a freshman who did not take a break. Also, similar analysis shows the same trend is also true for sophomores, juniors and seniors. However, the difference in dropout and graduation rates for seniors who take a break and those who do not take a break is very minimal (i.e. just 0.2%). This was expected considering the reasons why a student may take a break as outlined in the beginning of Section 4.

5 MODELING THE ENROLLMENT OF NEW STUDENTS

In this section, we consider incorporating the enrollment of new incoming students into our model. New incoming students are those entering ETSU for the first time irrespective of the class level they starts from, whether freshman, sophomore, junior or seniors. Although we have 18 semesters of data, we only considered new students enrolled starting from spring 2009 through Spring 2017. We do not have data on new students enrolled in fall 2008, hence it was excluded. In Matlab, we track the student id; if the id appears for the first time in a particular semester, the student is counted as newly enrolled. There are a total of 33,913 new students enrolled at ETSU for all 17 semesters in our data set. For all 8 fall semesters in the data, 28,017 were enrolled and 5,896 were enrolled for all 9 spring semesters. On average, approximately 3,503 and 656 new students were enrolled for each fall and each spring semester respectively. We can have a clear visualization of the total number of new students enrolled in ETSU from spring 2009 through spring 2017 from Figure 15 below.

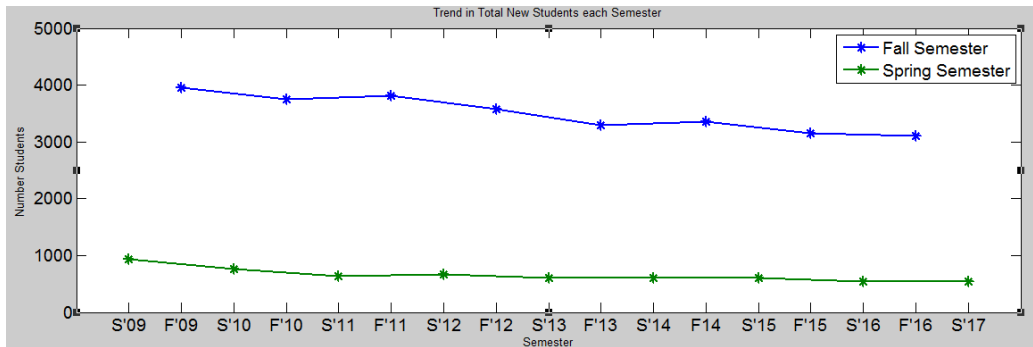


Figure 15: The Total New Student Enrolled

In Figure 15, we have the total new students per semester. For instance, in

Spring 2009 and Fall 2009 there are about 1,000 and 4,000 new students enrolled respectively. It can be seen from Figure 16 that the majority of new students are freshmen, followed by juniors, sophomores and seniors. Out of the total number of new students enrolled in our data, slightly over 20,000 are freshmen, slightly less than 4,000 are sophomores, about 6,000 are juniors and about 3,800 are seniors representing 60.2%, 11.6%, 16.9% and 11.4 respectively. For fall semesters slightly over 18,000 (64.8%) freshmen, slightly less than 3,000 (10.3%) sophomores, slightly more than 4,200 (15.2%) juniors and about 2,700 (9.7%) are enrolled. For Spring semesters, about 2,300 (38.4%) freshmen, about 1,050 (17.8%) sophomores, about 1,450 (24.6%) juniors and about 1,150 (19.4%) senior are enrolled.

We can see that there is a continuous decline in the trend for the total number of new students for each fall semester at ETSU. For the trend of new enrollment for each spring semester, the decrease in the number of new students enrolled from Spring 2009 to Spring 2011 is higher than for the following semesters.

It is necessary to look at the statistics of newly enrolled students for the various class levels to have a clear picture of what is happening. In Figure 16, we show the total number of new students and percentages for each semester. We can see that more new students are enrolled in fall semester than spring. There is an overall decrease in the total number of freshmen. However, the percentage of new freshmen students enrolled each fall is slightly increasing. Also, there is somewhat stability in the percentage of new students enrolled as sophomores, juniors and seniors in fall semester. Therefore, although the total number of new students is decreasing, the percentage enrolled into each class level is fairly stable. Similar features can be seen

for juniors and seniors in spring semester except new freshmen which rise in spring 2009, attained slight stability from spring 2010 to spring 16 and then start falling.

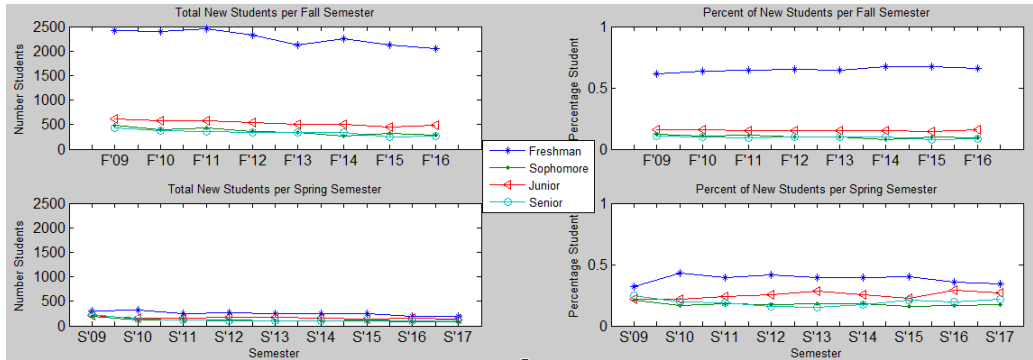


Figure 16: The Total New Student Enrolled By Class Level

In Figure 15, we approximate a regression equation for the total enrollment of new students for the past five years for fall semester which is given by $\hat{y} = -106x + 3618.5$ where y is the number of new students and x is the number of semesters, with $x = 0$ indicating Fall 2012. The slope is given by -106, this means the total number of new students enrolled each fall decreases on average by about 106 students per semester. For spring semester, the regression equation is giving by $\hat{y} = -22.3x + 646.1.3$. The slope for spring semester is -22.3, i.e. the total number of new students enrolled in spring semester decreases by about 22 students per semester. We use this equation when approximating how many students will enroll each semester in the simulations performed in the next section.

6 SIMULATION OF THE MODEL

In this section, we simulate the model to test the accuracy of the discrete-time Markov chain model in making predictions about student enrollment at ETSU. In the simulation, we look at the past five years (thus, from spring 2012 to spring 2017) and compare the trends. The simulation algorithm is given by the following:

Step 1: Initialize the number of student for the various ten states (we use the exact number of students in spring 2012).

Step 2: Initialize probability transition matrices P_{fall} and P_{spring} .

Step 3: Initialize the number of years or steps of the simulation ($n = 5$ years)

Step 4: Set $i = 1$ to $2n$ (indicating two semester per year), do the following:

Step 4a: Set P equal to the appropriate transition matrix depending on the semester (fall to spring or spring to fall).

Step 4b: Initialize the number of students in all states for the next semester to zero.

Step 4c: Estimate the number of new students for the semester using the appropriate regression equation for either fall or spring semester.

Step 4d: For each class of students, do the following:

Step 4d.i: Determine the transitions for the class using the appropriate column (P_j) of the transition matrix corresponding to the given class j .

Step 4d.ii: Generate a vector r of uniformly distributed numbers between 0 and 1 with length equal to the total number of students of the given class.

Step 4d.iii: Determine the number of students who transition to each state by determining the number of random numbers in r which are less than the cumulative sum of probabilities in P_j from Step 4d.i

Step 4d.iv: Add the total number of students for each class which come from all other states plus the new incoming students in each class. Label this vector v_i for semester i .

Figures 17-20 display the output of the simulations for five runs. Note that each run will be slightly different due to the vector of uniformly distributed random numbers from Step Step 4d.ii in the simulation algorithm. Figures 17 and 18 represents the total number of studentS enrollment at ETSU and their class levels respectively. Figures 19 and 20 show the total number of new students enrollment and their class levels respectively. From the figures, we compare the estimated number of student to the exact numbers given in the data and analyze the differences. If the plot of the estimated number of students equates or intersects the plot of the exact number of students, we assume there is a fairly accurate prediction. If the plot of the exact number of students is below the plot of the estimated number of student, we assume the number of students is overestimated and vice versa.

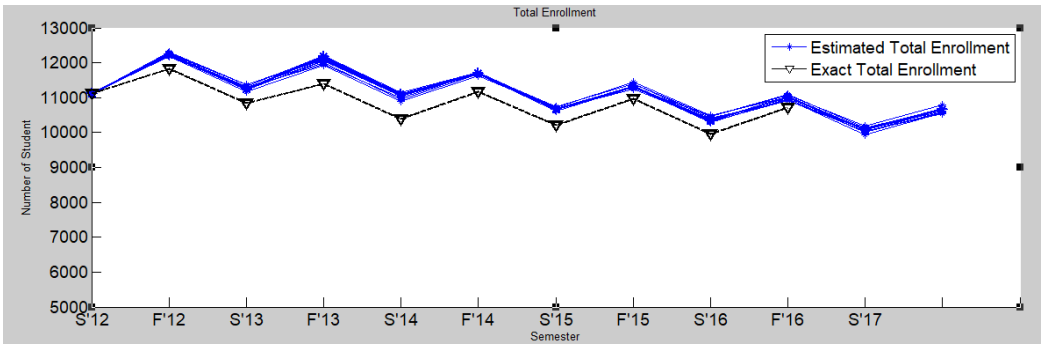


Figure 17: Simulation of Total Student Enrolled

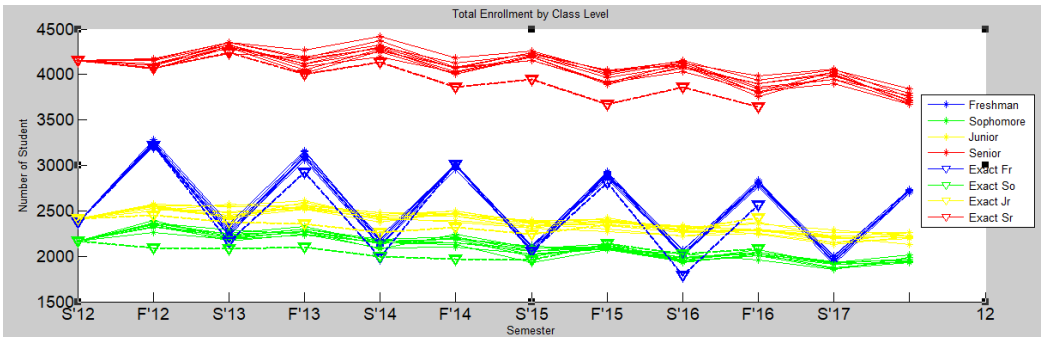


Figure 18: Simulation of Total Student Enrolled By Class Level

In Figure 17, the total number of students is overestimated by the model. From Spring 2012, the estimated total number of students enrolled rise above the exact total indicated by the data. Although both trends are decreasing, the total number of students is overestimated by approximately 10%. For instance, in fall 2012 we estimated about 12500 total students compared to about 11500 indicated by the data. It is also interesting to examine the behavior of the total number of students when divided according to class level as shown in Figure 18. In Figure 18, the initial behavior of the plot shows that there is somewhat accuracy in the prediction of the total number of seniors until Fall 2013. Thereafter, there is an overestimation of

seniors as the estimated number continues to spread apart from the exact number of seniors. This is most likely error in the estimated graduation rate for seniors. There is fairly good prediction for freshmen from spring 2012 to spring 2013 and from spring 2014 to fall 2015 since the total number of freshmen for both plots are fairly similar. For both sophomores and juniors we can see initially that there is an overestimation of the total number of students from spring 2012 to spring 2015. From Fall 2015 the prediction seems fairly accurate for both sophomores and juniors.

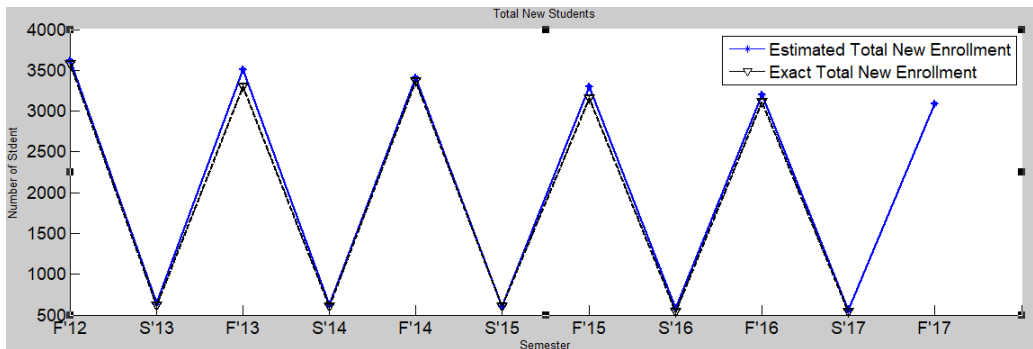


Figure 19: Simulation of Total New Student Enrolled

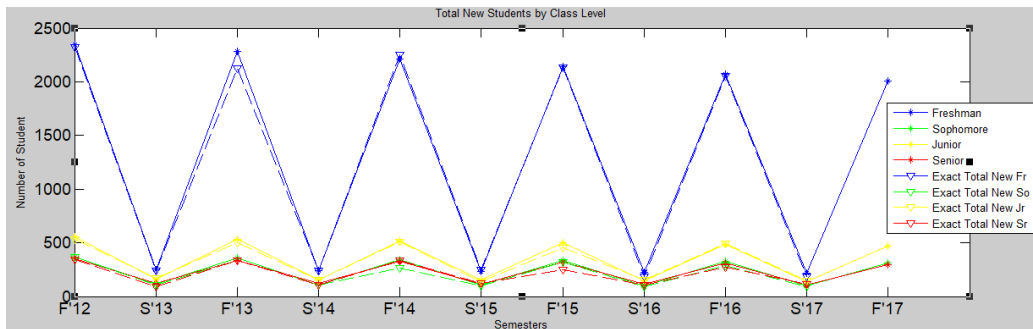


Figure 20: Simulation of Total New Student Enrolled by Class Level

In Figures 19 and 20, we simulate the total number of new student enrollment. We can see that there is a better prediction of the total number of new enrollment in

Figures 19 and 20 compared to the total number of students in Figures 17 and 18. The prediction of the enrollment of the total number of new student for the various class levels (Figure 20) is quite accurate except in a few cases such as freshmen for fall 2013 which can be seen to be slightly overestimated. Also, there is a slight overestimation in fall 2015 for juniors and seniors and in fall 2014 for sophomores.

7 CONCLUSION AND FUTURE WORK

In conclusion, in this thesis, we have used ETSU data from Fall 2008 to Spring 2017 to create a discrete-time Markov chain to model for undergraduate enrollment at ETSU. Initial analysis of the data indicates that more students are enrolled in fall semesters than spring semesters except seniors which has the opposite trend. On the contrary, more students dropout after spring semester as compared to fall semester. We used the data to build probability transition matrices to analyze how students transition through the various class levels. We concluded that the individual fall matrices are significantly not different from one another. Similar conclusions were made for spring matrices. By contrast, fall and spring matrices are shown to be different when using the matrix norm. Furthermore, the hiatus state was introduced to reduce the error in our model due to misclassification of dropouts. We determined that seniors are the most likely to take a break, representing 34.4% of the total number of students who took a break at some point. Also, when students take a break after fall semester (i.e. take spring semester off), they spend an average of one semester off and spend an average of two semesters off when taking a break after spring semester, i.e. if a student does not return for fall semester, the student is more likely to take an entire year off. We combined the hiatus states with the class level states to form a general probability transition matrix for fall and spring semester and then analyze potential trends using these probability transition matrices. In five years, it was predicted that there is a 26.0% probability of a freshman graduating and 58.5% probability of freshman dropping out. There is 82.0% of seniors graduating and 14.9% probability of seniors dropping out in five years. In the long-term, students

who take a break are seen to have a higher probability of graduating compared to those who do not take a break.

The inclusion of the hiatus state into our Markov chain model was necessary, and it increased the graduation rate of the number of freshmen after five years to 26% compared to the 21% predicted by Helbert's model [10]. The actual rate of graduation of a freshman after five years is 35% as indicated by the data. This means there are still discrepancies associated with our model. The simulation of the model shows it is fairly close to making accurate predictions about the total number of new students enrolled but overestimates the total number of students entirely. Hence, the reasons for the inaccurate prediction by the model is not only associated with misclassification of students but there are other factors which need to be considered in future work using the model. We hypothesize that in the future it might be necessary to reexamine how the probabilities are generated for the matrices.

BIBLIOGRAPHY

- [1] Doug Lederman. *The Enrollment Slowdown* [Online]. February 28, 2014. Available from: <https://www.insidehighered.com/news/2014/02/28/us-projects-college-enrollment-grow-14-through-2022>
- [2] Heather Long. *College enrollment is dropping. Bad sign?* [Online]. May 20, 2016: 12:00 PM ET. Available from: <http://money.cnn.com/2016/05/20/news/economy/college-enrollment-down/>
- [3] Tami Luhby, Tal Yellin and Caroline Matthews. *Just how much better off are college grads anyway?* Source: Bureau of Labor Statistics data for 2015; Published April 5, 2016. <http://money.cnn.com/infographic/economy/college-degree-earnings/index.html?iid=EL>
- [4] Kate Gibson / Moneywatch /May 15, 2015, 5:10 AM. *Why is college enrollment declining?*. <http://www.cbsnews.com/news/why-is-college-enrollment-declining/>
- [5] Jordan Friedman, Editor. Sept. 22, 2016, at 9:00 a.m. *10 Universities With the Most Undergraduate Students*. <https://www.usnews.com/education/best-colleges/the-short-list-college/articles/2016-09-22/10-universities-with-the-most-undergraduate-students>
- [6] Lalit Garg, Sally McClean, Brian Meenan and Peter Millard. 23 October 2009, Springer Science+Business Media, LLC 2009 *A non-homogeneous discrete time*

Markov model for admission scheduling and resource planning in a cost or capacity constrained healthcare system

- [7] T. W. Anderson and Leo A. Goodman. 07-02-2017 15:22 UTC. *Statistical Inference about Markov Chains*. Source: The Annals of Mathematical Statistics, Vol. 28, No. 1 (Mar., 1957), pp. 89-110. Published by: Institute of Mathematical Statistics. <http://www.jstor.org/stable/2237025>

- [8] Theodore Charitos, Peter R.de Waal and LindaC.van der Gaag *Computing short-interval transition matrices of a discrete-time Markov chain from partially observed data*. Statistics in Medicine. Statist. Med. 2008; 27:905921. Published online 20 June, 2007 in Wiley InterScience (www.interscience.wiley.com) DOI:10.1002/sim.2970

- [9] Polly Wainwright, Department of Mathematical Sciences Department of Computer and Informational Sciences Indiana University, South Bend. November 16, 2007. *AN Enrollment Retention Study Using a Markov Model for a Regional State University Campus in Transition*

- [10] Helbert, Zachery T., *Modeling Enrollment at a Regional University using a Discrete-Time Markov Chain* (2015). Undergraduate honor thesis. Paper 281. www.dc.etsu.edu/honors/281

- [11] Mike Bowler, Contributor — Aug. 19, 2009, at 4:00 p.m. *Dropouts Loom Large for Schools*. <https://www.usnews.com/education/articles/2009/08/19/dropouts-loom-large-for-schools>.

- [12] Terri Williams on July 8, 2015 AT 3:25 PM. *Freshmen Students Are The Most Likely to Drop Out of College.* <https://www.goodcall.com/news/why-freshman-are-the-most-likely-to-drop-out-of-college-01421>
- [13] Jordan Weissmann, MAR 29, 2012. *Why Do So Many Americans Drop Out of College?// How America's higher education system became one big dropout factory.* <https://www.theatlantic.com/business/archive/2012/03/why-do-so-many-americans-drop-out-of-college/255226/>
- [14] Julie Mack, jmack1@mlive.com on September 06, 2010. Report: *For college freshmen, first year is one of upheaval.* http://www.mlive.com/news/kalamazoo/index.ssf/2010/09/for_college_freshmen_first_yea.html
- [15] TNJ Staff, The Network Journal Thursday, July 16, 2015. *Should You Take a Break From School? 3 Reasons Why You May Need To.* <http://www.tnj.com/career/career-advice/should-you-take-break-school-3-reasons-why-you-may-need-to>
- [16] Eric Hoover July 14, 2011. *Recession Reshaped College Enrollment Patterns, but the Sky Didn't Fall.* <http://www.chronicle.com/article/Recession-Reshaped-College/128223>.
- [17] Ana Ley Las Vegas Review-Journal, August 29, 2015: *UNLV rebounding after post-recession enrollment slump.* <https://www.reviewjournal.com/news/education/unlv-rebounding-after-post-recession-enrollment-slump/>.

- [18] Jordan Weissmann, Nov 2014. *Americas Awful College Dropout Rates, in Four Charts*. http://www.slate.com/blogs/moneybox/2014/11/19/u_s_college_dropouts_rates_explained_in_4_charts.html.
- [19] Huang, Cheng-Chi, “Non-homogeneous Markov chains and their applications” (1977), Iowa State University. <http://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=8613&context=rtd>

VITA

LOHUWA MAMUDU

- Education: M.S. Mathematical Sciences,
East Tennessee State university, 2017
- B.S. Mathematics with Economics
University for Development Studies (UDS), 2014
- Professional Experience: Introduction to Probability and Statistics, T.A.
Department of Mathematics and Statistics,ETSU
Mathematics and Statistics Tutor
Center for Academic Achievement,ETSU
Jan. 2016 - Aug. 2017
- Professional Development: Statistical and Mathematical:
SAS, R, SPSS, Minitab, Matlab
- Programming Languages:
Microsoft Office Suite:
MS Access , Word, Excel, PowerPoint, Publisher,Outlook