

Life After High School

A Look at Factors that Impact the Completion of Post-Secondary Applications

By

Daniel Frederick Crites

A Thesis submitted in Partial Fulfillment

of the Requirements for the Degree

Master of Science in Mathematics

West Texas A&M University

Canyon, Texas

July 16, 2017

ABSTRACT

The purpose of this study is to examine the important factors that can lead a high school graduate to submit an application to a post-secondary institution. There has been a recent increase in the national high school graduation rate and yet a decrease in college enrollment of high school graduates. By identifying these influential factors, schools can see which students will likely need assistance in pursuing a post-secondary education.

This study reviews previous research to discuss what others have discovered in the area, as well as performs a logistic regression analysis to determine these factors based on a new dataset. This research will also generate a model to predict the likelihood of a student applying to a post-secondary institution, providing clues as to which factors may be more substantial.

Approved:

Thesis Committee Chairman

Date

Thesis Committee Member

Date

Thesis Committee Member

Date

Department Head

Date

Academic College Dean

Date

Graduate School Dean

Date

TABLE OF CONTENTS

Chapter	Page
INTRODUCTION	1
RESEARCH QUESTION	3
KEY TERMS	3
LITERATURE REVIEW	4
PARENTS	4
SOCIOECONOMIC STATUS	5
ACT/SAT EXAMINATIONS	5
PRIOR ACADEMIC PREPARATION/PERFORMANCE	6
COLLEGE ACCESS FOR THE HIGH PLAINS GRANT PROGRAM	8
METHODOLOGY	10
DATA AND COLLECTION PROCESS	10
DEPENDENT VARIABLE	12
INDEPENDENT VARIABLES	12
LOGISTIC REGRESSION	15
ANALYTIC DESIGN	16
ANALYSIS RESULTS	18
MODEL FOR COMPLETION OF POST-SECONDARY APPLICATION	25
PERFECT VARIABLE PREDICTORS	29
PERFECT INTERACTION PREDICTORS	30
CONCLUSIONS	31
LIMITATIONS	32
FUTURE RESEARCH	33
REFERENCES	34

LIST OF TABLES

Table	Page
DESCRIPTIONS OF NOMINAL PREDICTORS	14
DESCRIPTIONS OF DUMMY VARIABLES	15

LIST OF FIGURES

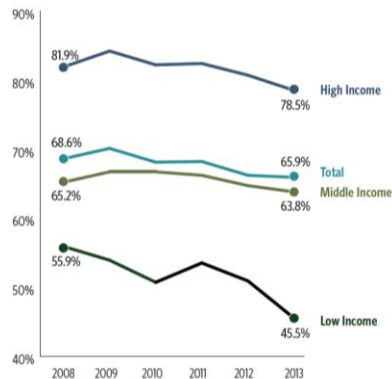
Figure	Page
FIGURE 1: COLLEGE ENROLLMENT OF HIGH SCHOOL GRADUATES	2
FIGURE 2: SINGLE LOGIT RESULTS AGAINST DEPENDENT VARIABLE APPLIED	18
FIGURE 3: PERFECT SUCCESSES.....	21
FIGURE 4: FIRST LOGISTIC REGRESSION RUN	22
FIGURE 5: REGRESSION AFTER REMOVAL OF FEMALE	23
FIGURE 6: REGRESSION AFTER REMOVAL OF GPA AND FEMALE	24
FIGURE 7: GOODNESS-OF-FIT TEST	26
FIGURE 8: GRADUATION PLAN COMPARISONS	28

CHAPTER I

INTRODUCTION

Higher education is one of the more talked about and debated topics in the United States. In today's modern society, many entry-level positions meant to help jumpstart the careers of the young adults require a bachelor's degree. Some of these positions even ask for a graduate degree. There is an abundance of resources for high school students when it comes to post-secondary preparation materials. Access to tutoring services, boot camps, college nights, test prep manuals, etc. are available across the country. According to the National Center for Education Statistics (NCES), the national high school graduation rate is roughly 83%, an all-time high for the United States. However, the US Census Bureau reported that college enrollment of high school graduates has dropped from 69% in 2008 to 66% in 2013 (See Figure 1).

Figure 1: College Enrollment of High School Graduates



SOURCE: U.S. Department of Commerce, Census Bureau, Current Population Survey (CPS), October, 2008 through 2013.
NOTES: Recent high school completers refers to individuals ages 18 to 24 who graduated from high school or completed a GED® test during the calendar year. Low income refers to the bottom 20 percent of all family incomes, high income refers to the top 20 percent of all family incomes, and middle income refers to the 60 percent in between.

While there are programs and supports in place for students to obtain their diplomas, programs designed to move students on to the next step in starting a post-secondary education are lacking effectiveness. This, combined with other factors at play, keep students from continuing their education beyond the secondary level. Some of these substantial issues are beyond the students' control, such as socioeconomic status and parental education levels. Other factors are in the students' control, such as academic performance. This study will look at previous research as well as examine new data to discover possible factors to determine when a student may need assistance finding or completing applications for post-secondary education. The data collected for this study comes from the researcher's participation in the College Access for the High Plains grant program as a "college access coach". This program is designed to assist Texas Panhandle high schools in ensuring their graduates enroll in post-secondary institutions. More information on this program is found in later chapters.

Research Question

There is a vast amount of literature on factors contributing to post-secondary enrollment. These will be presented in the literature review. However, a vital step in the process of post-secondary enrollment is applying to these institutions. The purpose of this study is to look at possible influential factors for high school seniors in their successful completion of post-secondary applications. To determine any such influences, consider the question:

In what ways do demographic, academic, and socioeconomic factors increase or decrease the likelihood that a high school senior will submit an application to a post-secondary institution?

Key Terms

Post-secondary institution: 4-yr Universities, 2-yr community colleges, technical/vocational schools, and the military are considered post-secondary institutions in this study.

Dichotomous: A categorical variable with two possible categories.

Logistic regression: A statistical method that takes one or more independent variables and determines the likelihood of the dependent variable occurring.

P-value: A statistic used to determine whether a variable is significant enough to be included in the regression model. A low p-value ($< .05$) indicates that the variable does have a significant impact on the model.

CHAPTER II

LITERATURE REVIEW

There has been a plethora of research on post-secondary education, especially enrollment. Factors that determine whether a high school student chooses to go to college and whether they enroll for school have been studied, as in the cases of Hossler & Stage (1992) and St. John (1991).

Parents

In many of these studies, it is found that parents play an integral role in the college process. When the parents are more knowledgeable about post-secondary topics such as college and financial aid, the students are more likely to attend a post-secondary institution (Engberg and Wolniak 2010; Plank and Jordan 2001). Also, when students and parents maintain discussions of matters relating to college, there is an increase in student enrollment (An 2010; Plank and Jordan 2001). However, when parents show concern over college expenses and financial aid, students are less likely to attend college (Engberg and Wolniak 2010). It would seem that high school students rely on the

knowledge and opinions of their parents to help them make decisions about their post-secondary plans.

Socioeconomic Status

Another major factor of influence in a student's decision to attend a post-secondary institution is socioeconomic status. Parents of middle- and upper-class families tend to be more involved in their student's post-secondary plans by seeking out help, visiting colleges, and consistently talking to their children about college (Lareau and Weininger 2008; McDonough 1994). However, lower-class as well as minority and first generation parents have limited knowledge of post-secondary education and therefore are not as involved (Ceja 2006; Lareau and Weininger 2008). Parents of upper- and middle-class families are more involved because they tend to understand the post-secondary education process (McDonough 1994). On the other hand, parents of lower-class families tend to be unaware of the process, which can be complicated at times, and even sometimes discourage their student from applying to more than one institution (Ceja 2006). To state it generally, "as socioeconomic status increases, students are significantly more likely to attend a 2- or 4-year college versus not enrolling at any college" (Engberg and Wolniak 2010, p. 143)

ACT/SAT Examinations

Besides influential factors outside of school, there are several important factors inside the school as well. One road block to acceptance into a post-secondary institution is taking the ACT or SAT examination. It is required by most 2- and 4-year colleges. Plank

and Jordan (2001) found that just by taking the test(s), there is an increase in the chance of attending an institution.

Prior Academic Preparation/Performance

How a student performs academically during high school also plays an important role.

Taking Advanced Placement courses has a positive impact on students attending 4-year universities (Engberg and Wolniak 2010). Success in regular high school courses are also important because research has found that increases in overall grade point average (GPA) leads to an increase in the likelihood of attending a post-secondary institution (Engberg and Wolniak 2010; Manski and Wise 1983).

Finally, the high school itself can be a significant factor in the attendance of a post-secondary institution by a high school graduate. Whether a school has a college-going culture can influence a student's enrollment after high school. In schools with a higher college-going culture, it is assumed that students will attend college, so there is information, tools, and resources at the school counselor's disposal. The counselors then give their students focused advice and attention on the college process, typically starting as early as 9th or 10th grade (Engberg and Gilbert 2014). On the other side of the spectrum, high schools with a low college-going culture typically give much less attention, if any, to counseling students on college (Engberg and Gilbert 2014; McDonough 1997). In these kind of schools, the counselor's role focuses more on disciplinary issues, scheduling classes, and even counseling students on career choices and prep (McDonough 1997). If there is any information given to students regarding

college, it is usually very late in their high school career and it usually contains basic and generic information that may result in students looking at 2-year colleges, if any (McDonough 1997).

There are many different areas of life that can positively or negatively affect a high school student's post-secondary plans. Some of these factors may seem more (or less) significant than others, but each one is a stepping stone on the path from high school graduation to post-secondary education. Not only that, but each student is unique in that the priorities of these influential factors vary from person to person. Even one stone missing can lead to an unstable path and that path may end up deteriorating and the student may never reach post-secondary education.

CHAPTER III

COLLEGE ACCESS FOR THE HIGH PLAINS GRANT PROGRAM

This study uses the information gathered through the College Access for the High Plains (CAHP) program. This program, working in conjunction with West Texas A&M University's offices of Academic Affairs and Enrollment Management, began in 2014 in order to develop a high quality, multi strategy college access program for the Texas Panhandle region. According to the Texas Education Agency's 2012 Academic Excellence Indicator System (AEIS) Report of Canyon ISD, roughly 66% of seniors took the ACT and/or SAT examination, which indicates interest in post-secondary education. However only 47% of seniors enrolled in post-secondary institutions within 12 months of graduation. CAHP seeks to drastically raise these numbers, especially the disconnect between showing interest in attending college and enrolling, by providing several services to the high school, parents, and students.

The services performed by the program include ACT test preparation classes, tutoring services, parent information nights, surveying of students' post-secondary knowledge and aspirations, and aiding high school seniors in the post-secondary education process. The surveys were conducted one at a time and in person for high school seniors and

facilitated by English classes for the freshmen, sophomores, and juniors. More information on the content of the survey can be found in the next chapter.

The focus of this study is on the surveying and aiding of high school seniors. As part of the CAHP program, “college access coaches” meet with seniors to discuss post-secondary plans and aid in achieving those plans, if needed. During the initial meeting, information is obtained and recorded through the survey, the specifics of which are discussed in the subsequent chapter. If a senior is found to be behind in various deadlines, then one or more follow-up meetings are scheduled to ensure that the student is making progress in achieving their plans.

CHAPTER IV

METHODOLOGY

Data and Collection Process

The data for this study was obtained from the 2014 – 2015 graduating class of Canyon High School. There were 230 seniors that took part in the CAHP program in the spring of 2015.

Demographic and academic information on each subject was obtained through high school transcripts and recorded into a database. This information included:

- Gender – coded as Male or Female
- Race – coded as American Indian/Alaskan Native, Asian, African American, Native Hawaiian, and Caucasian
- Ethnicity – coded as Hispanic/Latino and Not Hispanic/Latino
- Graduation track – coded as Recommended, Distinguished, and Foundation
- GPA – on a 100 scale
- Class rank
- Number of Dual Credit courses taken, and

- Number of Advanced Placement courses taken

Class rank is a quantitative variable based solely on GPA. A higher GPA means a higher class ranking. Since GPA is already included as a possible predictor, including class rank would be superfluous. As such, this variable was not included in the study. Race was another variable not included in the study, as explained later in the limitations section. Besides class rank and race, the remainder of the variables collected from the transcripts were used in this study.

Further information was obtained on CAHP participants through survey responses; all of which is used in the study. This information included:

- Whether the student is on free or reduced lunch – coded as Yes or No
- Whether the student would be a first-generation college student – coded as Yes or No
- What type of post-secondary path the student decided to pursue – coded as 2-year college, 4-year university, and Other (workforce, military, vocational/technical, and undecided)
- Whether the student completed the Free Application for Federal Student Aid (FAFSA) – coded as Yes or No
- Whether the student has started applying for scholarships– coded as Yes or No, and

- The number of times the student had to be followed up with by a “college access coach”

Dependent Variable

The dependent variable used in this study is if the student has completed and submitted at least one application for a post-secondary institution. There are two categories: the student completed at least one application; the student did not complete an application.

Independent Variables

There are four ratio level variables and eight nominal level variables used as predictors.

The ratio level variables, as well as their label in STATA, are:

- The number of dual credit courses taken throughout high school (DC)
- The number of Advance Placement courses taken throughout high school (AP)
- The student’s GPA (GPA), and
- The numbers of times the student had to follow up with the “college coach” (NUMFOLLOWUP)

The nominal level variables, as well as their label in STATA, are:

- Gender (FEMALE)
- Ethnicity (HISPANIC)
- Whether the student is on Free or Reduced Lunch (LUNCH)

- Whether the student would be first-generation (FIRSTGEN)
- Whether the student completed the FAFSA (FAFSACOMPLETED)
- Whether the student completed at least one application for scholarships (SCHOLARSHIPS)
- The student's post-secondary plans (DECISION)
- The student's graduation track (GRADTYPE)

Regression analysis requires all nominal variables to be dichotomous. Dummy coding was used to reformat all nominal variables to a dichotomous format. Consider for example the student's post-secondary decision. There were three possible categories for a student's post-secondary plans: attend a 2-year community college, attend a 4-year University, or Other. To make this dichotomous, two dummy variables are created. Dummy variable 'CC' has values '0' and '1', where '0' represents not going to a 2-year college and '1' represents going to a 2-year college. Dummy variable 'UNI' also has values '0' and '1', with '0' representing not going to a 4-year University and '1' representing going to a 4-year University. If both 'CC' and 'UNI' have values of 0 for a particular student, then that student falls into the category of 'Other'. Table I gives the descriptions of all the original nominal variables while Table II illustrates how the dummy variables resulted in the dichotomous property that was needed.

Table 1
Descriptions of Nominal Predictors

Gender	0 - Male 1 - Female
Ethnicity	0 - Not Hispanic or Latino 1 - Hispanic or Latino
Graduation Type	0 - Recommended 1 - Distinguished 2 - Foundation
Free/Reduced Lunch	0 - Does not receive free or reduced lunch 1 - Does receive free or reduced lunch 2 - Don't know
First-Generation Student	0 - Would not be a first-generation student 1 - Would be a first-generation student 2 - Don't know
Post-secondary Decision	0 - 2-year college 1 - 4-year university 2 - Other
FAFSA Completed	0 - FAFSA is not complete 1 - FAFSA is complete
Applied for Scholarships	0 - Have not applied for scholarships 1 - Has applied for scholarships

Table II
Descriptions of Dummy Variables

REC	0 - Not on Recommended plan 1 - On Recommended plan If REC and DIS are both 0 -- On Foundation plan	DIS	0 - Not on Distinguished plan 1 - On Distinguished plan
YL	0 - No free/reduced lunch 1 - Free/reduced lunch If YL and NL are both 0 -- Student doesn't know	NL	0 - Free/reduced lunch 1 - No free/reduced lunch
FG	0 - Not first-generation 1 - First-generation If FG and NFG are both 0 -- Student doesn't know	NFG	0 - First-generation 1 - Not first-generation
CC	0 - Not going to 2-yr college 1 - Going to 2-yr college If CC and UNI are both 0 -- Student chooses Other	UNI	0 - Not going to 4-yr University 1 - Going to 4-yr University

Logistic Regression

In this study, the dependent variable is a dichotomous nominal-level variable. The predictors are made up of dichotomous nominal-level variables as well as continuous and discrete ratio-level variables. A binary logistic regression model is most appropriate to use in this case. This particular model is used because of the types of variables involved (McDonald 2014) as well as the fact that this study is looking at the likelihood of a student completing applications.

The general form of a logistic regression model is $\sigma(t) = \frac{1}{1+e^{-t}}$ where t represents the linear function made up of the predictor variables x_i . This function has the form $t =$

$\beta_0 + \beta_1x_1 + \dots + \beta_nx_n$ for n independent variables with β_0 the constant term and β_i the regression coefficients for each particular predictor. A positive coefficient means that for β_i increase in the predictor, the likelihood of the dependent variable occurring increases. A negative coefficient means that for β_i increase in the predictor, the likelihood of the dependent variable occurring decreases.

For this study, the model used is

$$\text{logit}(\mathbb{E}[Y|x_1, \dots, x_n]) = \beta_0 + \beta_1x_1 + \dots + \beta_nx_n$$

where $\text{logit}(\mathbb{E}[Y|x_1, \dots, x_n])$ is the expected log of the odds that a student will complete a college application.

Analytic Design

The first step in this analysis is to check for any missing data values and adjust if needed.

For this particular dataset, there were no missing values among the variables so no adjustments are necessary.

The statistical software used for this study is STATA. It uses logistic regression function *logit*. The logistic regression in this study uses a backwards elimination method. The regression is first built by looking at a single logit model with the dependent variable and each independent variable. Using the p-value, possible significant variables are determined for the initial full model regression function. The first logistic regression model is built using all significant predictor variables. STATA's program runs through multiple iterations of the model, trying to maximize the log likelihood since logistic

regression uses the maximum likelihood. The program will keep running iterations until the difference between the previous model and the next model is small enough to say that it has converged. At that point STATA stops running iterations and displays the results of the regression. A backwards elimination process is then used to obtain the final model. In this process, the least significant predictor is removed and the regression is run again. If the coefficients of the remaining predictors have not changed by more than roughly twenty percent, then the removal was necessary. However, if there is a drastic change in coefficients then there is a significant interaction and the removed predictor is put back into the model. This procedure is repeated until all insignificant predictors have been removed from the model.

The final step is to check and see if there are any interactions between predictors that may be significant. If there are, then they are to be included as well. Once this is completed, the final regression model is built using any significant predictors and significant interactions.

Once a final model is developed, it is tested for the goodness of fit using STATA's built in goodness of fit test with the command *estat gof*.

CHAPTER V

ANALYSIS RESULTS

The first step in the analysis is running a single logit model with each independent variable. Figure 2 shows the results for each possible predictor.

Figure 2: Single Logit Results against Dependent Variable APPLIED

Independent Variable	Coefficient	Std. Error	z	P> z 	[95% Conf. Interval]	
FEMALE	0.84601	0.4984827	1.70	0.09	-0.130998	1.823018
HISPANIC	-0.5499164	0.5119	-1.07	0.283	-1.553222	0.4533892
DC	0.5160375	0.1774814	2.91	0.004	0.1681803	0.8638946
AP	0.3379548	0.2703066	1.25	0.211	-0.1918363	0.8677459
GPA	0.0809887	0.0369953	2.19	0.029	0.0084792	0.1534982
NUMFOLLOWUP	0.1640342	0.2589805	0.63	0.526	-0.3435583	0.6716266
YL	-0.4809735	0.5438338	-0.88	0.376	-1.546868	0.5849211

Independent Variable	Coefficient	Std. Error	z	P> z 	[95% Conf. Interval]	
NL	-0.2735831	0.5317251	-0.51	0.607	-1.315745	0.7685789
FG	-0.5909562	0.4719721	-1.25	0.211	-1.516005	0.3340921
NFG	0.2968565	0.4684089	0.63	0.526	-0.6212082	1.214921
FAFSA	Perfect Success					
SCHOLARSHIPS	Perfect Success					
REC	0.9092895	0.4650303	1.96	0.051	-0.0021531	1.820732
DIS	0.7726935	0.6413146	1.20	0.228	-0.4842601	2.029647
CC	Perfect Success					
UNI	3.14075	1.032906	3.04	0.002	1.116292	5.165209
RANK	0.0081615	0.0032833	-2.49	0.013	-0.0145966	-0.0017264

These single logit models revealed several pieces of important information. Using the p-values, predictors used in the next phase of the regression are identified. These variables and their p-values are as follows:

- FEMALE - Gender (0.09)
- DC - Number of Dual Credit Classes (0.004)
- GPA (0.029)
- REC - Recommended Graduation Plan (0.051)
- UNI - Deciding on a 4-yr University (0.002)

While the variables REC and FEMALE have p-values over 0.05, they are still close enough to put into the initial regression model.

The other important information in these models is the discovery of perfect successes in the variables SCHOLARSHIPS (has applied for scholarships), FAFSA (completed the FAFSA), and CC (decided on a community college). This means that there is no instance where a student completed at least one scholarship application and did not complete at least one post-secondary application. Similarly, every student who completed the FAFSA also completed at least one post-secondary application. Likewise, every student who decided on attending a community college had completed at least one post-secondary application as well. Figure 3 shows this in the form of two-way tables. A further discussion of this is in the results section.

Figure 3: Perfect Successes

Community College	Completed Application for Post-Secondary		Total
	0	1	
0	22	162	184
1	0	46	46
Total	22	208	230

FAFSA	Completed Application for Post-Secondary		Total
	0	1	
0	22	122	144
1	0	86	86
Total	22	208	230

Scholarships	Completed Application for Post-Secondary		Total
	0	1	
0	22	117	139
1	0	91	91
Total	22	208	230

Using the independent variables deemed significant by the individual logit models, an initial regression is run. Figure 4 displays the results from this.

Figure 4: First Logistic Regression Run

APPLIED	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
FEMALE	0.6209569	0.5362254	1.16	0.247	-0.4300257	1.671939
DC	0.3901292	0.224712	1.74	0.083	-0.0502982	0.8305566
GPA	-0.0609313	0.0507891	-1.2	0.23	-0.1604761	0.0386135
REC	0.9765884	0.5069142	1.93	0.054	-0.0169452	1.970122
UNI	2.838233	1.083074	2.62	0.009	0.7154462	4.961019
_cons	5.76034	4.232682	1.36	0.174	-2.535564	14.05624

The p-value for the initial run of the regression model is less than 0.05, so the model is statistically significant. The least significant variable in the model is FEMALE (gender), with a p-value of 0.247. It is removed and the regression run again, as can be seen in Figure 5.

Figure 5: Regression after removal of FEMALE

APPLIED	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
DC	0.413326	0.2234916	1.85	0.064	-0.0247095	0.8513615
GPA	-0.0560521	0.051078	-1.1	0.272	-0.1561632	0.0440589
REC	1.048314	0.50594	2.07	0.038	0.0566903	2.039939
UNI	2.797518	1.079509	2.59	0.01	0.68172	4.913317
_cons	5.509723	4.262047	1.29	0.196	-2.843736	13.86318

Since the coefficients of the remaining variables did not change by a significant amount, the process is repeated with the variable GPA removed. This can be seen in Figure 6.

Figure 6: Regression after removal of GPA and FEMALE

APPLIED	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
DC	0.2784867	0.177473	1.57	0.117	-0.069354	0.6263274
REC	0.9220439	0.4893163	1.88	0.06	-0.0369984	1.881086
UNI	2.648409	1.070096	2.47	0.013	0.5510596	4.745757
_cons	0.8657244	0.3457556	2.5	0.012	0.188056	1.543393

Looking at the coefficient for the number of Dual Credit courses (DC), it has changed significantly, incurring a roughly 33% reduction. This indicates that there is an interaction involving GPA and another predictor that is significant, even if GPA itself is not significant. Because of this, GPA needs to remain in the regression model.

At this point there are four main predictors left in the regression model: the number of dual credit courses taken, the student's GPA, the student being on the recommended graduation track, and whether the student decided to go to a 4-year university. The p-values for the model as well as the four predictors (save for GPA) are all less than 0.05, indicating a statistical significance.

The next step is to check for any significant interactions between the predictors. There were four main interactions that resulted in perfect success:

- When the student took at least three dual credit courses and decided to either attend a 4-year university or not,
- When the student took at least three dual credit courses and either is or is not on the recommended graduation track,
- When the student took at least three dual credit courses combined with their GPA, and
- When the student is on the recommended graduation track and decided on attending a 4-year university.

The remaining interactions were not statistically significant and therefore not included in the model.

Model for Completion of Post-Secondary Application

The model for the student completing and submitting at least one application for a post-secondary institution can be found by the following:

$$\begin{aligned} \text{logit}(E[Y|x_1, \dots, x_n]) &= \text{The expected log odds of completing an application} \\ &= 0.4133(x_1) - 0.0561(x_2) + 1.0483(x_3) + 2.7975(x_4) + 5.5097 \end{aligned}$$

where x_1, \dots, x_4 are the values for the following predictors:

- $x_1 = \text{number of dual credit courses taken}$
- $x_2 = \text{GPA}$
- $x_3 = \text{Recommended graduation track}$

- x_4 = whether the student decided on a 4 year university

The final step in the regression is to check the model to make sure the data fits well.

Figure 7 shows the results of the goodness-of-fit test. The p-value for the test is 0.8314, meaning the regression model cannot be rejected and hence is a good model to use for the data.

Figure 7: Goodness-of-fit Test

Logistic model for APPLIED, goodness-of-fit test

number of observations	230
number of covariate patterns	230
Pearson chi2(225)	204.62
Prob > chi2	0.8314

When looking at the independent variables, some may be more predictable than others. First, a student can earn both college and high school credit for completing a dual credit course with the benefit of being able to take the course at their high school. More dual credit classes can be interpreted as more free college credit the student can receive prior to attending college. If a student takes multiple dual credit courses, it can be assumed that they are motivated to pursue a post-secondary education. Otherwise they would likely not take this type of course.

Second, if a student decides on attending a 4-year university, then there is most likely motivation behind their decision for various reasons. They may have a career in mind that requires a high-level degree. They may want to attend the same college their parent(s) did. They may even just want to follow their close friends. Also, by deciding on a 4-year university they most likely know that an application is needed to attend the institution.

Third, it is significant if the student is on the recommended graduation plan. Comparing the recommended plan to the minimum plan, as can be seen in Figure 8, several aspects appear to be influential in the student's education career. First, students only need three math credits and Algebra II is not a requirement on the minimum plan. However, with the recommended plan students need four math credits, including Algebra II. Second, minimum plan students do not require a foreign language, while recommended students need two credits (equivalent to a full year) of foreign language. Third, the minimum student needs only two science credits while the recommended student needs four science credits. By having to take more science, math, and language classes the student may feel more motivated to continue their education beyond high school. They will be taking science and math all four years of high school while minimum students can be finished with those requirements by their junior year, going two years without math and science.

Figure 8: Graduation plan comparisons

Discipline	*MHSP	*RHSP	*DAP
English Language Arts	Four credits: <ul style="list-style-type: none"> English I English II English III English IV or approved alternate course 	Four credits: <ul style="list-style-type: none"> English I English II English III English IV 	Four credits: <ul style="list-style-type: none"> English I English II English III English IV
Mathematics	Three credits: <ul style="list-style-type: none"> Algebra I Geometry SBOE approved math course 	Four credits: <ul style="list-style-type: none"> Algebra I Algebra II Geometry An additional math credit 	Four credits: <ul style="list-style-type: none"> Algebra I Algebra II Geometry An additional math credit
Science	Two credits: <ul style="list-style-type: none"> Biology IPC or Chemistry and Physics (one of the two serves as an academic elective) 	Four credits: <ul style="list-style-type: none"> Biology Chemistry Physics An additional science credit 	Four credits: <ul style="list-style-type: none"> Biology Chemistry Physics An additional science credit
Social Studies	Three credits: <ul style="list-style-type: none"> U.S. History (one credit) U.S. Government (one half credit) Economics (one half credit) World History (one credit) or World Geography (one credit) 	Four credits: <ul style="list-style-type: none"> U.S. History (one credit) U.S. Government (one half credit) Economics (one half credit) World History (one credit) World Geography (one credit) 	Four credits: <ul style="list-style-type: none"> U.S. History (one credit) U.S. Government (one half credit) Economics (one half credit) World History (one credit) World Geography (one credit)
Physical Education	One credit	One credit	One credit
Languages Other Than English	None	Two credits in the same language	Three credits in the same language
Fine Arts	One credit	One credit	One credit
Speech	One-half credit from either of the following: <ul style="list-style-type: none"> Communication Applications Professional Communications (CTE) 	One-half credit from either of the following: <ul style="list-style-type: none"> Communication Applications Professional Communications (CTE) 	One-half credit from either of the following: <ul style="list-style-type: none"> Communication Applications Professional Communications (CTE)
Electives	Seven and one half credits (one must be an academic elective)	Five and one-half credits	Four and one-half credits
Total Credits	22	26	26

The final predictor is the student's GPA. With a p-value of 0.272, the variable itself is not statistically significant. If it were significant, then it would logically seem that it should have a positive coefficient, not the negative one that it does in fact have. When this variable interacts with the number of Dual Credit courses, then there is a significant interaction, which is discussed below.

Perfect variable predictors

At the beginning of the regression process, there were three variables that had perfect success. One of those variables is the student deciding on attending a community college. In this instance, every high school senior who made the decision to go to community college had submitted the application for their desired school. One reason for this could be the perception that community colleges are inherently “easier” than 4-year universities. If the student thinks that the school and/or the application process will be easier then they may be more inclined to complete the application, even though it may be similar, if not identical, to a 4-year university application. There is also a societal consensus that for students who struggle in school or financially, it could be better to go to a 2-year college first before transferring to a 4-year university.

The second perfectly successful variable was when the student had submitted the FAFSA. There could be a few reasons for this. Many students only look at a financial aid form once they know they are attending some sort of college, which they know after they complete an application. Another possible explanation could be that the student has a strong support system when it comes to college plans. A sibling or parent or someone close may have gone through the process and is coaching the student through it. In this case, having this support system could greatly help the student in submitting various applications and forms such as the FAFSA.

The third variable with perfect success was when the student had submitted scholarship applications. The reasons for the FAFSA variable could easily be applied to this one as

well. A student will most likely not seek out and apply for scholarships until they know they need to help pay for college. They will not know if they need to until they know they are going to college, which requires the completion of a post-secondary application.

Perfect interaction predictors

Recall the logistic regression model had only the four significant variables. Interaction terms were then considered and four of the possible interaction terms also predicted success perfectly. The first is when the student is on the recommended graduation track and decides to attend a 4-year university. The increase in core courses and the desire to go to a university could show the level of motivation needed to apply to that university.

The three other interactions all deal with Dual Credit courses. Any student who took three or more Dual Credit courses applied to a post-secondary institution regardless of being on the recommended plan or deciding on a 4-year university. Also, when the student took three or more Dual Credit courses in combination with their GPA ensured the submission of a post-secondary application. These are interesting interactions, indicating that by willingly taking more than two dual credit courses the student is more motivated to their schoolwork and committed to attending college. Successfully completing three Dual Credit courses means nine free college credit hours, almost an entire semester. The student knows that will be able to start college ahead of what is considered “normal”.

CHAPTER VI

CONCLUSIONS

The results of this study revealed several important factors in determining if a high school student will complete a post-secondary application. Deciding on applying to a community college, completing the FAFSA, and completing scholarship applications yielded perfect successes in completing at least one college application, suggesting that these are very influential factors in making sure a post-secondary application is completed. Deciding on a 4-year university and being on the recommended graduation plan also would seem to have a positive impact on the likelihood of applying to a post-secondary institution. Finally, taking three or more dual credit courses in high school can be a major motivator for a student to complete a post-secondary application, especially since that is almost an entire college semester that the student has earned free of charge. These factors can be used by high schools to raise the rate of graduating high school students going to college immediately after high school graduation. Secondary institutions should consider promoting and encouraging enrollment in dual credit courses, talking to students early about plans after high school to get them to decide before graduating, and offering services or instruction to help students complete the

FAFSA and start filling out scholarship applications.

Limitations

This study has several limitations. The data was collected by taking part in an already established grant program at West Texas A&M University. This in itself may not be a limitation, but it creates a few of them.

First, the grant program initially collaborated with one high school in the Texas panhandle region and data was collected on one senior class. This has led to a small and localized dataset. Because of the dataset, student race had to be dropped from the analysis. A variable that has many groups, like race, would cause separation issues in the logistic regression model. Also, when the regression was run the first time, several variables had to be dropped due to resulting in perfect success or failure, which is elaborated on in the analytic design section.

Second, the main purpose of the grant program was to help high school seniors prepare for post-secondary education and was not accomplished for the soul purpose of this study. As a result, some of the variables collected for the grant program were not relevant for this study and some variables that could have been relevant for this study were not for the grant program and so were not collected.

One limitation that does not come from the grant program is collecting part of the data through surveying the students. Survey questions have long been documented as having bias issues. When people are interviewed, they often want to give what they think are

answers the interviewer wants them or expects them to give (Clark and Schober 1992; Terry and Hogg 1996). People tend to exaggerate in order to give the appearance of being better, participating more, etc. as well (Holbrook and Krosnick 2010; Barker et al. 1994; Brenner and DeLamater 2016; Klesges et al. 1990). In this study, since the students knew they were meeting with 'college coaches', they may have given responses that they thought would seem favorable. Also, there is a definite societal norm of going to college. With the abundant amount of scholarships, college prep resources, tutoring services, etc. in today's world, there is an emphasis placed on a college education. High school students see all of this and may give false answers on the survey in order to 'fit in' with what is considered normal.

Future Research

The first obvious expansion of this research is to incorporate other schools into the study. One place to start would be with one whole district or region. This could help a district or region determine if there are students who may need assistance in their post-secondary education. The next level would be to branch out into the rest of the Texas panhandle area. A study at this level would let schools know the factors that impacts the post-secondary plans of an average student from the panhandle. By expanding the scope of this study, many of the limitations and issues discussed may be lessened or even become non-issues. It would also be beneficial to duplicate this study since the state graduation plans have changed quite significantly.

CHAPTER VII

REFERENCES

- An, B. P. (2010). The relations between race, family characteristics, and where students apply to college. *Social Science Research*, 39, 310-323.
- Barker, Kathleen, Lynda Fong, Samara Grossman, Colin Quin, and Rachel Reid. 1994. "Comparison of Self-Reported Recycling Attitudes and Behaviors with Actual Behavior." *Psychological Reports* 75:571-77
- Brenner, Philip S. and John DeLamater. 2016. "Measurement Directiveness as a Cause of Response Bias: Evidence from Two Survey Experiments." *Sociological Methods & Research* 45:348-371.
- Ceja, M. (2006). Understanding the Role of Parents and Siblings as Information Sources in the College Choice Process of Chicana Students. *Journal of College Student Development*, (1), 87.
- Clark, Herbert H., and Michael F. Schober. 1992. "Asking Questions and Influencing Answers." Pp 15-48 in *Questions about Questions: Inquiries into the Cognitive Bases of Surveys*, edited by J. M. Tanur. New York: Russell Sage.
- Engberg, M. E., & Gilbert, A. J. (2014). The counseling opportunity structure: Examining correlates of four-year college-going rates. *Research in Higher Education*. 55(3), 219-244.
- Engberg, M. E., & Wolniak, G. C. (2010). Examining the Effects of High School Contexts on Post-secondary Enrollment. *Research in Higher Education*, (2). 132.
- Holbrook, Allyson L., and Jon A. Krosnick. 2010. "Social Desirability Bias in Voter Turnout Reports: Tests Using the Item Count Technique." *Public Opinion Quarterly* 74:37-67.

- Hosmer, David W., Lemeshow, Stanley, and Sturdivant, Rodney X. 2013. *Applied Logistic Regression*. Jon Wiley and Sons, Inc.
- Hossler, D., & Stage, F. (1992). Family and High School Experience Influences on the Post-secondary Educational Plans of Ninth-Grade Students. *American Educational Research Journal*, 29(2), 425-451.
- Klesges, Robert C., Linda H. Eck, Michael W. Mellon, William Fulliton, Grant W. Somes, and Cindy L. Hanson. 1990. "The Accuracy of Self-Reports of Physical Activity." *Medicine and Science in Sports and Exercise* 22:690–97.
- Lareau, A., & Conley, D. (2008). Chapter 4: Class and the Transition to Adulthood. In , *Social Class : How Does it Work?* (p. 118). New York: Russell Sage Foundation.
- Manski, C.F., & Wise, D.A. (1983). *College choice in America*. Cambridge, MA: Harvard University Press.
- McDonald, J.H. 2014. *Handbook of Biological Statistics*, 3rd ed. Sparky House Publishing, Baltimore, Maryland
- McDonough, P. M. (1994). Buying and Selling Higher Education: The Social Construction of the College Applicant. *The Journal of Higher Education*, (4). 427.
- McDonough, P. M. (1997). *Choosing Colleges: How social class and schools structure opportunity*. Albany: State University of New York Press.
- Plank, S. B., & Jordan, W. J. (2001). Effects of Information, Guidance, and Actions on Post-secondary Destinations: A Study of Talent Loss. *American Educational Research Journal*, (4). 947.
- Terry, Deborah J., and Michael A. Hogg. 1996. "Group Norms and the Attitude-Behavior Relationship: A Role for Group Identification." *Personality and Social Psychology Bulletin* 22:776–93.
- U.S. Department of Education, National Center for Education Statistics. (2009)Table 1. Public high school 4-year adjusted cohort graduation rate (ACGR), by race/ethnicity and selected demographics for the United States, the 50 states, and the District of Columbia: School year 2014–15. Retrieved from https://nces.ed.gov/ccd/tables/ACGR_RE_and_characteristics_2014-15.asp