

Factors that Predict the Probability of Student Success at Missouri Southern State University

By

Ryan Prisock

In

Economics/Finance

A Thesis

Presented to the

Honors Program of

Missouri Southern State University

May 2013

Factors that Predict the Probability of Student Success at Missouri Southern State University

Abstract

This study examined the success of rates of students in the 2011 freshman class at Missouri Southern State University, with success being defined as their first-year college GPA or if they returned for a sophomore year. The study used ACT score, high school GPA, high school rank, age, age², gender, a composite admissions decision criteria, and major status as independent variables. High school GPA, age, and age² were significant predictors of College GPA. High school rank, high school GPA, and the composite admissions decision criteria variable were significant predictors of retention to a sophomore year. The study discusses the implications of these results, as well as some of the limitations in the analysis.

Keywords: High school GPA, high school rank, ACT score, gender, age, major, student success, retention, college GPA, freshman

INTRODUCTION

Through the opening years of the 21st Century, the benefits and costs of a college education has been a topic of both increasing frequency and controversy. Higher education has long been viewed as a pathway to increased economic and social benefit, both for the individual student and society at large. However, these benefits do not come without cost, costs which have gained increasing weight in light of continual government deficits and rising student debt levels. Due to this it has become increasingly important to show the benefits of a college education are worth its investment, and that this investment is made to the benefit of both the student and society.

While the benefits of a college education are multifold, the economic advantages are often those most frequently cited as justification for its costs. This is justified when one sees the economic advantage that a college education conveys, as a college education is beneficial in securing employment, higher salaries, and greater benefits (College Board Advocacy and Policy Center, 2010, p.3-4). These advantages have only increased in importance in the wake of the Great Recession of 2007-2009, during which unemployment skyrocketed, peaking at 10.0% in the months directly after the recession (Bureau of Labor Statistics [BLS], 2012, p. 2). Though the recession affected employment across the board, those with higher education suffered less severe and sustained effects (Pew Charitable Trusts, Jan 2013, p.11).

The benefits of college education should not be limited to those of an economic nature, however. Though receiving less focus, societal benefits are also a primary factor in evaluating the worth of a college education. Controlling for income, the “non-market” benefits of higher education include better health, greater longevity, healthier and better educated children, and greater happiness (McMahon, Feb 2009, p.119). Society also benefits from education, with

greater national income growth, economic impact from universities, greater participation in politics, lower crime rates, healthier environments, and increased human rights. (McMahon, Feb 2009, p. 183-229,). These benefits are often more difficult to quantify, yet they are of indispensable value in establishing the worth of investment in education.

Despite the established benefits of higher education, concerns about the cost of about college education and access have continued to mount in recent years. In a 2009 study by The National Center for Public Policy and Higher Education and Public Agenda, 83% of Americans believe that students have to borrow too much to pay for their education. 69% believe that qualified people don't have the opportunity to get a higher education. 65% believe that college prices are growing faster than other things. All of these numbers were higher than previous year marks (The National Center for Public Policy and Higher Education and Public Agenda, Feb 2010). These concerns are compounded by states nationwide spending 28% less per student in higher education from 2008 to 2013, while tuition has increased nationwide 27% in the same time period. (Center on Budget and Policy Priorities, March 2013). Student debt has also skyrocketed, both in dollars and as a national concern. As of 2012, student loan debt stood at \$966 billion, tripling between 2004 and 2012. 17% of borrowers are over 90 days in delinquent, with 30-49 year olds having the highest delinquency rates. These numbers have continued to trend upwards at exponential rates (Lee, Feb 2013)

In light of the importance of higher education and its rising costs, it has become more important than ever to ensure that that the investment made in higher education continues to provide the value that both students and society expect. The belief in college education to provide economic stability and mobility continues to draw students, with enrollment increasing 37% between 2000 and 2010 (National Center for Education Statistics, May 2011). In addition to

the increase in enrollment, it is worth focusing on where this increase is coming. In the 2000 to 2010 timeframe, female enrollment outpaced male enrollment in undergraduate institutions at 39% to 35%, respectively. Also, college student bodies increasingly becoming older; Enrollment of students 25 and older has increased 8 percentage points higher than students under 25, at 42% to 34%. This trend is expected to become even more pronounced in the next ten years. (National Center for Education Statistics, May 2011).

Considering the rising costs and changing student demographics, it has become more important than ever to ensure the benefits of society are capable of meeting both its costs and the needs of those who seek education. To do this, it is prudent to examine factors that determine who is admitted to college, what factors are used to predict their success, and whether or not these factors give a picture that is both accurate and fair to the students they effect.

In looking at the 2011 freshman class at Missouri Southern State University, this study seeks to examine a few of the factors that are either used in the admissions process or aid in describing the entering student body as a means of determining the probability of the student achieving success at Missouri Southern. First-year college grade point average and retention to sophomore year are used as the parameters which describe student success. This study hopes to quantify the effect that some common factors among entering college students have on these parameters, and to provoke discussion on these relevance of these parameters and how they can be used to the benefit of future students, Missouri Southern State University, and the Joplin community.

LITERATURE REVIEW

In determining which factors to examine for predicting the probability of student success, a natural starting point is to examine which quantitative factors Missouri Southern State

University uses in its admissions process. There are three quantitative factors which are used in the admissions criteria, of which the student need only meet one to secure admittance: a student's American College Testing Composite Score, more commonly known as the ACT Composite Score, must be 21 or higher, high school grade point average (GPA) must be 2.25 or higher, or the student's high school class rank must be in the top 50% of their class. Meeting one of these criteria meets the admissions criteria, though it is still possible to gain admittance on review.

Of these three quantitative factors, ACT score is one that receives prominent attention. The ACT is very popular nationwide, with 47% of all high school graduates taking the test in 2010. This exhibits an increase in the popularity of the ACT, as 40% of high school graduates took the test in 2006. More importantly, however, is the ACT's popularity in MSSU's geographic region. Missouri and the surrounding states of Iowa, Illinois, Kentucky, Tennessee, Arkansas, Oklahoma, Kansas, and Nebraska all have over 60% of their high school graduates take the ACT, with Illinois, Kentucky, Tennessee, and Arkansas having more than 80% of high school graduates taking the ACT exam. (ACT Inc., 2010)

Given the number of residents in Missouri and neighboring states who take the ACT exam, the increasing number of high school graduates who take the exam, and Missouri Southern State University's policy of making it one of their three quantitative factors for admissions, the ACT Composite score is a natural choice to examine as a probability predictor of student success at Missouri Southern. This choice is aided by the fact that extensive research has already been done on the ACT and its predictive abilities.

In previous studies done on the ability of the ACT score to predict the probability of student success, success has primarily been defined in two ways: college GPA and retention. The selection of these two variables is an intuitive one, as each measures an integral component of

success. Retention shows the steady intent and progression towards the student's degree, while college GPA serves to measure the degree of success that the student enjoys. Studies have thus utilized these outcomes as dependent variables upon the ACT in the hopes of determining its effects.

In both college GPA and retention, the ACT has shown to be a predictor of college success. One study measured the ACT college readiness benchmarks, which are 18, 22, and 24 for English, Math, and Science, respectively, upon college GPA and retention. Using analysis of college grade data, the probability of attaining success, as defined by a grade of B or higher in an introductory course, was found to be 50 percent at the defined benchmarks (Allen and Sconing, 2005). The probability of a grade of C or higher in an introductory course was 75-80 percent if the ACT college readiness benchmarks were met. Also, if these benchmarks were met, the probability of a freshman returning for his sophomore year was 83 percent (ACT, Inc., 2005).

Even when not using the established benchmarks, the ACT composite score has been shown to predict the probability of student success. In a sampling of 7,377 first-year college students, a logistical regression analysis was used to show ACT scores have a positive correlation with both college GPA and college retention, with correlation of two-year college GPA and retention being 34.1 percent and 14.5 percent, respectively (House, 1994). These results have been supported by additional empirical analysis as well; in a separate study, ACT scores accounted for 26.5 percent of the variance in college GPAs (Myers and Phyles, 1992). Additionally, while ACT is not always the best predictor of the probability of student success in grades or retention, it is still positive predictor of college GPA, particularly for the first-year of college (Noble and Saywer, 2002). Based on these studies, it can be reasonably inferred that

ACT has a positive predictive ability on the probability of student success. However, this is not to imply that it is the best predictor.

Additional factors used in predicting collegiate success are high school class rank and GPA. These are the other two factors that Missouri Southern State University uses as a quantitative measure for admissions. These factors have shown a positive correlation with college success as well, being able to beat even ACT scores in predictive ability. One instance of this correlation found correlations with high school rank of 47.2 percent and 22.7 percent to 2-year college GPA and retention, respectively (House, 1994). High GPA has also shown better predictability for certain ranges of college GPA, being able to more effectively predict success of college GPA levels 2.00-3.00, but losing predictability at higher levels (Noble and Sawyer, 2002).

High school GPA maintains predictive ability without a grade range as well, accounting for 18% of the variance in college GPA for the first year (Myers and Phyles, 1992). Further support of high school GPA predictability can be found in another study, where logistical regression analysis established correlations for both ACT score and high GPA to college GPA. In 3 of the 4 years sampled, high school GPA showed a stronger correlation than ACT scores in predicting college success (Bontakeoe, 1992). Though strong correlations can be established with both ACT and high school predictors, it should not be ignored that these factors might be also be affected by other variables.

In several of the studies cited, the predictive ability of ACT and high school GPA and rank were affected by the gender of the student. In the majority of studies, the predictive ability of both ACT scores, high school GPA, and rank on college success was stronger for women than it was for men (House, 1994.; Myers and Phyles, 1992.; Bontakeoe, 1992.). However, this should

not be viewed as a foregone conclusion, as one study could find no discernible difference in the correlation between gender and the ACT or high school GPA or rank's ability to predict college success (ACT, Inc., 2008). In light of these conflicted studies, it can be inferred that the only way to be clear of the factors effects on a given institution is through looking at the specific institution and the other variables that have an impact on predicting college success. While those previously mentioned have received wide focus, others have remained less researched.

There are two other variables that have not been used heavily in conjunction with a model that includes ACT scores or high school GPA and rank. These two variables are age and major.

Jacobs (2002) demonstrated that being over 25 years old in college puts one at a disadvantage when seeking their degree, as the women in the study were often enrolled part-time and had to deal with other competing factors for their time and attention (Jacobs, 2002). Missouri Southern State University does not require an ACT for those 21 years of age or older, and it would be interesting to see if at this age cutoff there is any effect on the probability of college success.

The second variable is whether the student is a declared or undeclared major. It has been shown that changing majors in the freshman year can make the likelihood of that freshman staying increase by a factor of 4 (Synco, 2012). Considering this, someone coming in as an undeclared versus declared major could have a profound effect on their retention, and thus the predictability of the retention factors. In both age and major status, it remains to be seen how their predictive ability would effect and be affected by the other predictive factors previously mentioned.

As shown through the literature, the topic of factors predicting college success is not a new one. However, there are several intriguing things about the topic. College success, though sometimes defined using retention, is primarily defined by college GPA in the majority of the studies. Also, the factors predictability of college success and effect upon each other, speaking of

ACT scores, high school GPA and rank, and gender, varies depending on institution. In addition to this point, the effect of the determination of major or age upon entering has barely been considered in the context of this subject. Thus, the conclusion can be made that there is still knowledge to be uncovered in order to provide a more comprehensive picture of college success prediction factors and their interactions.

METHODOLOGY

The analysis in this study was based off of the first year statistics of the 2011 entering freshman class at Missouri Southern State University. The data was collected as part of the admissions procedure, and was made available to the Plaster School of Business through the Center for Institutional Research at MSSU. A total of 718 students were a part of the freshman class, consisting of 358 females and 347 males, with 13 having unreported gender.

The study covered all the students of the freshman class which had complete data observations. ACT Composite score, age, gender, college GPA, high school GPA, high school rank, whether the student entered as a declared or an undeclared major, and whether or not the student returned for a senior year were the data points used in the analysis. Of the original 718 students, 623 had each of the data observations and were used in the study.

GPA REGRESSION ANALYSIS

In the first set of regression analyses, college GPA is specified as a function of high school rank, high school GPA, gender, being a declared or undeclared major, ACT Composite score, and age. Descriptive Statistics for the variables are presented in Tables 1. The descriptive

statistics are broken down into the male, female, and total number of students, with the p-value denoting statistical difference also listed.

COLLEGE GPA is a continuous interval variable that demonstrates both academic ability and effort in college. MSSU grades on 4.0 scale, with 0 being assigned for grades lower than a D, 1.0 for Ds, 2.0 for Cs, 3.0 for Bs, and 4.0 for As. Of the 623 students utilized in the analysis, 25% had a GPA of 1.95 or lower, 50% had a GPA of 2.86 or lower, and 75% had a GPA of 3.47 or lower. These percentiles varied by gender, with males faring poorer than females. Males had a mean college GPA of 2.40 and females had a mean college GPA of 2.69. The difference between these two means was statistically significant, a p-score of .000974, giving a greater than 99% confidence level that the two means are statistically different from one another.

Missouri Southern State University evaluates progress using a student's GPA and number of credit hours completed. In order to remain in good academic standing, a student must retain a GPA of 1.5 or higher for the first 22 credit hours completed. As more hours are taken, the required GPA for good academic standing increases, with a 2.0 GPA required in order to graduate with either an associate's degree or a bachelor's degree. Of the 2011 freshman class, 69.5% of males and 78.9% of females had first year GPAs of 2.0 or higher at Missouri Southern State University. The difference in the proportion of male and female students who finish is statistically significant, with a p-value of less than .00001.

AGE is a continuous interval variable that indicates the age of the student when they first enrolled at Southern. Values range from a minimum of 16 to a maximum of 41, with the mean age being 18.57 and the median being 18. The expectation is that age will correlate positively with college GPA, as it can be used as an indicator of maturity or the perhaps signal greater seriousness with which one takes their studies.

A second regression analysis was done using AGE^2 as an additional variable. AGE^2 is also a continuous interval variable that indicates the inverse, parabolic relationship that age typically has (age is beneficially until a prime is reached, and then starts to work against the individual). The minimum, maximum, means, and medians are the same as the squares of the corresponding variables in AGE. The belief was that AGE^2 would correlate negatively with college GPA, as other factors not present in this study could manifest more frequently in older students.

GENDER is a dummy variable, with 1 being assigned if female and 0 being assigned if male. The 13 unreported students were eliminated from the study. The assumption made is that being female would positively correlate with college GPA. This is indicated by both the literature review and the descriptive statistics regarding GPA, with the mean female GPA being 2.69 compared to the mean male GPA of 2.40, though the gap in GPA between males and females closed as grade levels increased. As with the college GPAs, these figures proved to be statistically different from one another, with a p-value of .0000055, indicating a .00055% probability that the two GPAs are statistically equal to one another; we can be over 99% confident in the statistical difference of the two numbers.

The student's MAJOR was a dummy variable, with 0 being undeclared and 1 being declared. Of the 623 students which were used in the analysis, 498 (or 79.3%) had declared majors, with 129 (20.7%) of the students having undeclared or undecided majors. It was not known what effect if any this variable might have on college GPA.

ACT SCORE is a continuous interval variable that claims to measure the aptitude that a student has for college. The scale ranges from 0 – 36, with the national average being under 21; the average ACT score for MSSU students was also 21, and was so for both genders. The highest

admitted student's ACT SCORE was a 34 and the lowest admitted student ACT SCORE was a 10. Overall, 25% of the students were an 18 or lower and 75% of the students were a 24 or lower. As the literature review suggests, the belief was that higher ACT scores would correlate positively with higher college GPAs.

HS RANK is a continuous interval that ranges from 0 to 100 percentile. Higher values correspond with a higher placement amongst peers at students' high school institution. The value tells what percentage of a student's peers rank at or lower than the student i.e., a rank of 84 means that 84% of the class was equal to or below the student's performance. This variable was expected to positively correlate with College GPA, as students who exhibit consistent effort to achieve a higher ranking most like have socio-economic and individual attributes that transfer to their college career. The minimum high school rank was 5 and the maximum was 100, with a median of 66. 25% of the students were ranked 45 or lower and 75% of students were ranked 82 or higher.

HS GPA is a continuous interval that ranges from 0.0 to 4.45. Higher GPA values are indicative of greater academic achievement at the high school level. This variable was also expected to correlate highly to College GPA for the same reasons as high school rank; a student who performs well in high school most likely has transferable attributes, both socio-economic and individually, that will contribute favorably to their college success. The minimum HS GPA for females was 1.892 and 2.02 for males and the maximum HS GPA was 4.33 for females and 4.45 for males. The median HS GPA was 3.34, with 25% being less than 2.89 and 75% being more than 3.69.

The mean HS GPA was 3.18 and 3.36 for males and females, respectively. The difference between the two means was statistically significant, with p-value of 0.000011. This gives a

.0011% probability that the two numbers are statistically the same, establishing a confidence interval in well over 99% that the two numbers are statistically different.

COLLEGE GPA REGRESSION ANALYSIS RESULTS

The first college GPA regression analysis returned an adjusted R-squared value of .37210, indicating that the examined factors account for 37.21% of predicting the probability of college success as determined by College GPA. The model proved to be significant with an F score of 62.434 and an F significance of 1.96701E-60. This states that the model as specified is significant in explaining 37.21% of the variance in college GPA. The results of the regression analysis are summarized in Table 1.

The second college GPA regression analysis, which added AGE² returned an adjusted R-squared value of .37572, indicating that addition of AGE² increased the account of the variance by .362% to 37.572%. This model also proved to be significant with an F score of 54.478 with an F significance of 1.7202E-60. As stated previously, this describes the model as specified as being significant in explaining 37.572% of the variance in college GPA. The results of the second regression analysis are summarized in Table 3.

Results from the first regression analysis model (Table 2) indicate that College GPA is positively correlated with the AGE, GENDER (females are stronger than men), ACT SCORE, HS GPA, and HS RANK. The only negatively correlated value was MAJOR, with undeclared majors having higher GPAs than declared majors. However, only two of the variables were shown to be statistically significant at $\alpha = .01$: AGE and HS GPA. Using the p-values of .00182614 and 4.6792E-12 for AGE and HS GPA, respectively, the two variables are shown to have confidence levels over 99%.

Results from the second regression analysis model (Table 3) indicate that College GPA is positively correlated with the AGE, GENDER (females are stronger than men), ACT SCORE, HS GPA, and HS RANK factors, and negatively correlated with AGE² and MAJOR. As with the first analysis, both AGE and HS GPA were statistically significant at $\alpha = .01$ with p-values of 0.011412703 and 4.49718E-12, respectively. Additionally, AGE² was also statistically significant at $\alpha = .05$. AGE and HS GPA continue to have confidence levels of over 99%, with AGE² having a confidence level of over 95%. Using the coefficients and differential calculus, the optimal age for a student was found to be 31 years old. This was found by taking the derivative of the equation $GPA = \beta_1 * AGE^2 + \beta_2 * AGE$, setting it equal to zero, and solving for AGE, which equaled 31 as previously stated.

As mentioned, the results show that only two factors are statistically significant in predicting college GPA: HS GPA, AGE, and AGE². These factors help illustrate which individual characteristics might be most valuable in determining the probability collegiate success. As stated, AGE was expected to correlate positively due to a possible maturity or effort benefit that older students would have over younger students. The correlation proved true, though it cannot be concretely determined why in the scope of this study that age has the positive benefit upon college GPA that it does. It could be that those who are older appreciate the opportunity more and give better effort. It could also be true that those who are older are more inclined to view college more for its educational offering than for its social experiences; older students are in a different phase of life, being more likely to be married and/or have kids and as such may be more focused on home life than college social life. Maturity or increased skill in time management could also be a factor, as older students, having to already learn a work-life balance, are more readily able to transfer these skills than their younger counterparts. However,

as demonstrated above, after age 31, the effect started to reverse itself, which might be a sign that the complexity of factors associated with non-traditional students eventually increase in complexity to the point that they negate any of the speculated benefits of age (maturity, etc.).

HS GPA was also expected to correlate positively, which proved true in the regression analysis. While it is not in the scope of this study to suggest why positively related variables have the effect that they have upon college GPA, it can be speculated that success in both is a product of the same socio-economic and individual factors. Past research has suggested that this is indeed the case. In addition to middle school grades and standardized test scores from middle school, academic achievement, psychosocial, behavioral, demographic, and school-based factors have been shown to play a part in predicting the probability students' success in determining high school GPA (Allen et al, 2012). Other studies have pointed to the Big Five personality traits, with conscientiousness and openness both being characteristics that positively relate to academic achievement (Nofle and Robins, 2007). Similar factors as the ones in this study might similarly explain some of the correlation that HS GPA has, though without a more detailed study it cannot be stated for certain.

RETURN LOGISTIC REGRESSION ANALYSIS

In analyzing RETURN as the dependent variable, a logistic regression analysis was utilized due to the dependent variable being a binary variable RETURN from freshman year to sophomore year was specified as a function of HS RANK, HS GPA, GENDER, MAJOR, ACT SCORE, and AGE in the first analysis. The second included the AGE² variable. Descriptive Statistics for the variables are presented in Tables 1, with descriptive statistics being broken

down into the male, female, and total number of females, with the p-values denoting statistical difference also listed.

RETURN is a binary variable, with 1 indicating that the student returned for their sophomore year, and a 0 indicating that the student did not return. Of the 623 students who were included in the study, 387 (62.12%) returned for their sophomore year, with 216 of 331 females returning (65.26%) and 171 of 292 males returning (58.56%). The difference between these two proportions is significant, with a p-value of .0006, which corresponds to a confidence level over 99%.

As it had been in the linear regressions, AGE was expected to correlate positively with the probability of return for a second year, for many of the same factors mentioned previously (maturity, work ethic). ACT was expected to positively correlate as well, along with HS RANK and HS GPA. GENDER was expected to positively correlate with the probability of return, with females being more likely to enroll for a sophomore year. MAJOR status was also expected to positively correlate with the probability of return, with the belief that more determinant plans might signal a greater willingness pursue higher education.

A second logistical analysis was run with the added AGE² variable. This variable was added to accommodate for the inverse parabolic nature of that age demonstrates as it increases. AGE² was expected to negatively correlate with RETURN, as factors associated with age (family, financial independence, job, etc) grow in both magnitude and frequency.

RETURN LOGISTIC REGRESSION ANALYSIS RESULTS

The results from the first logistical regression analysis (Table 4) show that the majority of predictive factors behaved according to expectations. The probability of RETURN was

positively correlated with AGE, HS RANK, HS GPA, and GENDER (females have an advantage to men). Two of the variables, MAJOR and ACT SCORE, were negatively correlated. As with the regression analysis, however, only two of the variables were statistically significant: HS RANK and HS GPA. HS RANK has a p-value of .0871, signifying statistical significance at $\alpha = .10$, and HS GPA has a p-value of .0432 and a corresponding statistical significance at $\alpha = .05$. As it was one of the statistical significant variables in the linear regression analyses, AGE had a p-value of .1153, putting it just outside the range of statistical significance at $\alpha = .10$.

In the second logistical regression analysis (Table 5), the results do not significantly change. AGE, HS RANK, HS GPA, and GENDER still correlate positively, with MAJOR and ACT SCORE correlating negatively; AGE² was negatively correlated as well, in accordance with expectations. However, it was not statistically significant. HS RANK and HS GPA were also the only variables to be statistically significant, with the shifts in their p-values being slight enough that they remained significant at the $\alpha = .10$ and $\alpha = .05$ levels, respectively. The main difference in the logistical regression analysis after the addition of AGE² is that AGE received a large drop in significance, with its p-value moving from .1153 to .4362. This would seem to indicate that there is not a wide enough distribution in ages in order to fully observe the effect that AGE and AGE² has as it shifts from a beneficial to a negative factor as it increases in value.

These results show that only two factors are statistically significant in predicting the return from freshman to sophomore year: HS RANK and HS GPA, with AGE being worth mention due to its presence in the previous analyses. As these results are consistent with the previous ones, the hypothesized reasons for their predictive ability on COLLEGE GPA might also be applied in understanding their predictive ability on RETURN. It is worth noting,

however, that while these variables remained significant, their significance did decrease in predicting RETURN. This would seem to be in line with expectations, however, as the factors that might affect COLLEGE GPA during an academic year would not necessarily be the same or of the same magnitude as those that effect RETURN in between academic years.

ADC LOGISTIC REGRESSION ANALYSIS

In order to further analyze RETURN, a second set of logistic regression analyses was performed, with RETURN serving as the dependent variable. In this regression analysis, RETURN served as a dependent variable of AGE, GENDER, MAJOR, and ADC. ADC, which stands for admissions decision criteria, is a composite dummy variable which replaced ACT SCORE, HS RANK, and HS GPA. The analysis was run a second time, with AGE² being added to the independent variables. Descriptive Statistics for the AGE, GENDER, and MAJOR can be found in Table 1, and descriptive statistics for ADC can be found in Table 6.

As mentioned previously, Missouri Southern State University utilizes three quantitative variables, each with its own benchmark: an ACT SCORE of 21 or higher, a HS GPA of 2.25 or higher, and HS RANK of 50% or higher. Only one of these benchmarks is required to be met in order to gain admissions in to Missouri Southern State University.

Rather than examine the three variables separately, a composite value denoting how many of the benchmarks were reached was used to examine the admissions criteria as a whole. The admissions decision criteria variable, or ADC, is a dummy variable that denotes the degree of benchmarks the student has reached, with a 0 indicating no benchmarks were reached, and integers 1, 2, and 3 indicating the corresponding number of benchmarks which were reached. Of the 623 students in the sample, 14 (2.25%) met 0 of the 3 benchmarks, 127 (20.39%) met 1 out

of the 3 of the benchmarks, 190 (30.50%) met 2 out of the 3 benchmarks, and 292 (46.87%) met all three benchmarks. More detailed statistics are included in Table 6.

In light of the previous analyses, AGE, GENDER, and ADC were expected to correspond positively with RETURN, with the probability of continuing to a sophomore year rising as AGE, GENDER, and ADC increased. MAJOR was expected to negatively correlate given its previous predictive signs.

ADC Logistic Regression Results

The first ADC logistic regression analysis showed that AGE, GENDER, MAJOR, and ADC all had a positive relationship with RETURN. AGE, GENDER, and ADC meet expectations, with the only deviation being MAJOR's reversal, showing that those with declared majors are more likely to enroll for a second year. Of the four values, only ADC was statistically significant, with a p-value of less than 0.0001, showing statistical significance at $\alpha = .01$. These results are summarized in Table 7.

The second ADC logistic analysis showed nearly identical results, with AGE, GENDER, MAJOR, and ADC all having a positive predictive relationship upon RETURN. AGE² again had a negative predictive relationship with RETURN, keeping with previous observations. The statistical significance was not noticeably changed. AGE, GENDER, MAJOR, and AGE² were all statistically significant, with ADC being the only statistically significant variable. ADC maintained a p-value of less than .0001 and statistical significance at $\alpha = .01$. These results are summarized in Table 8.

The predictive ability and statistical significance of the ADC variable seems to that an increase in predictive ability is obtained when the admissions factors are used in conjunction

with each other. In the previous analysis, HS GPA was the primary determinant for college success, with HS RANK being the best secondary determinant. ACT SCORE was not statistically significant at all. The strong predictive ability of the ADC variable indicates its effectiveness as a composite, pulling the varying predictive factors of the three participatory variables into cohesion. This seems to validate both the ADC's use in college admissions, and the use of its component variables.

DISCUSSION

In light of the results of this study, there are several things of which to take note. The first observation that can be made is in defense of an open admissions policy. While Missouri Southern State University has admissions requirements, a student may still be admitted even if they do not meet one of the required benchmarks. Due to this, the number of students admitted who have a low probability of returning for a sophomore year or of achieving a satisfactory GPA will definitely be higher than if strict admissions standards were enforced. However, as was shown in the linear regression analysis, there are limitations to these predictors of success as well, with only high school GPA being a statistically significant for both retention and college GPA. As high school GPA is more of a measure of consistent effort than ACT, that same effort can aid in overcoming the disadvantages which might have led to a lower ACT score.

Another point the research raises is the potential benefit in waiting to enter college. AGE was another factor which consistently proved to be both positively correlated and very significant in predicting college success. Though it seems that the preparation for college is increasing shifting younger, it might be both more beneficial for the student and society at large if more time is taken to before enrollment to clarify the student's goals and college's roles in

filling them. This might especially be useful to students already at a higher risk to not achieve college success. This is not to trivialize the increased challenges that come with being an older student, as the research did suggest that the advantage brought about by a greater age began to reverse itself in time.

As high school GPA was the one consistent, statistically significant predictor of both success in college GPA and retention to a second year, a point illustrated by the research is the importance of the high school years in preparing the student to achieve success. As high school can be either a terminal or preparatory step in education, its importance cannot be understated. The research further illustrates that in the context of college success, with those who have developed or possess the characteristics necessary to do well in their high school studies have the advantage in continuing this success in college.

Perhaps the most important thing that this research illustrates, however, is the presence of a moral question. Though the admissions decision criteria was shown to be a positive, statistical predictor of the probability of a student returning for their sophomore year, over 20% of students were admitted after meeting 1 or less of the admission decision criteria. As has already been stated, open enrollment is not a bad policy, as it gives students the chance to achieve success based on their current efforts and abilities. However, if open enrollment is an encouraged policy, than it will naturally be the case that students of higher risk for failure will be taken. Thus, an obligation is created by the undertaking that all steps are taken to ensure that these students have the best possible chance available to them, despite any previous challenges to their success.

LIMITATIONS

In analyzing the results of this research, there are a few things which must be considered as possible limitations of its explanatory ability. The first of these is the data from which this analysis has been done. This research focuses exclusively on the 2011 Missouri Southern State University freshman class, which limits the data in three key ways.

The data is limited is by the research being constrained to Missouri Southern. As the literature review shows, similar analysis of predictors of college GPA and retention from freshman to sophomore year have been done at other institution, and the prevalence of these predictors in institutions across the country makes it appropriate to view the results of this and similar research in reference to each other. However, the data is specifically unique to Missouri Southern, and as such the results are as well. A different freshman class at a different institution would certainly yield at least slightly different results, the magnitude of which may or may not be statistically significant.

The results of this data analysis are focused purely on the data from the 2011 freshman class, and these results are most relevant in the context of first year college GPA and retention. The ability of this study to predict college GPA and retention beyond the freshman year is one that is primarily anecdotal due to no analysis being done with college GPA or retention beyond the first year. As such, this research is most relevant in a shorter time frame than a full college career.

As stated, the results of this analysis are based on the MSSU freshman class of 2011. However, a freshman class is not an accurate picture of a university's student body, as it does not factor in transfer students, students who are returning to finish their degree after hiatus, or

students who are returning for a second degree. The ability of this study to predict the probability of their college success is beyond the scope of this study to evaluate.

In addition to the limitations of the data set, there also some issues which must be considered in the analysis. In the linear regression analysis with college GPA as the dependent variable, the model is significant, performs for the most part as hypothesized, and has two statistically significant variables in the first analysis and three in the analysis which factors in age squared. However, the adjusted R-square is .3721 and .3758 for the two analyses, showing that the model can only explain about 37% of the variance in College GPA; over 62% of the variance cannot be explained by the model. Part of this may be due to the fact that none of the variables are purely socio-economic or behavioral in nature, but are rather indicators of the presence of such factors (Allen et al, 2012 and Nofle and Robin, 2007) This does not invalidate the results; it simply reinforces that the data be interpreted in light of a limited context.

Overall, both the linear regression model and logistic regression models are limited in that they only have two significant variables. While the independent variables' signs primarily behaved in accordance with expectations, the lack of statistical significance necessitates that the results of the regression analyses be reviewed with caution.

References

- ACT, Inc. (2005). ACT College Readiness Benchmarks, Retention, and First Year College GPA: What's the Connection? Retrieved July 1st, 2011, from ProQuest database
- ACT, Inc. (2010). The Condition of College and Career Readiness. Retrieved July 1st, 2011, from ProQuest database.
- ACT, Inc. (2008). What do we know about college success: Using ACT data to inform education issues. Retrieved July 1st, 2011, from ProQuest database
- Allen, Jeff and Sconing, Jim (August, 2005). Using ACT assessment scores to set benchmarks for college readiness. Retrieved July 1st, 2011, from ProQuest database
- Allen, Jeff, Casillas, Alex, Hanson, Mary Ann, Kuo, Yi-Lung, Robbins, Steve, and Schmeiser, Cynthia. (May, 2012). Predicting early academic failure in high school from prior academic achievement, psychological characteristics, and behavior. *Journal of Educational Psychology, Vol 104(2), May, 2012. pp. 407-420*
Retrieved April 8th, 2013, from EBSCO host database.
- Bontakoe, Jon F. (December, 1992) The ACT as a predictor of College Success at Trinity Christian College. The ACT as a Predictor of College Success at Trinity Christian College. Retrieved July 1st, 2011, from ProQuest database.

Bureau of Labor Statistics. (February, 2012). *The recession of 2007-2009*. Washington, D.C.

Center on Budget and Policy Priorities. (March, 2013). *Recent deep state higher Education cuts may harm students and the economy for years to come*. Washington, D.C. Leachman, Michael; Johnson, Ingrid; Oliff, Phil; and Palacios, Vincent.

Clements, William H. (August, 1969). Relationship of academic predictors to college success. Office of Institutional Research, Wisconsin State University. Retrieved July 1st, 2011, from ProQuest database.

College Board Advocacy and Policy Center. (2010). *Education Pays 2012: The benefits of higher education for individuals and society*. New York, NY: Baum, Sandy, Ma, Jennifer, and Paylea, Katherine.

House, J. Daniel (November, 1994). College grade outcomes and attrition: an exploratory study of noncognitive variables and academic background as predictors. Northern Illinois State University. Retrieved July 1st, 2011, from ProQuest database.

Jacobs, Jerry A., and King, Rosalind Berkowitz. (July, 2002). Age and College Completion: a life-history analysis of women aged 15-44. *American*

Sociological Association, vol. 75, no. 3. Retrieved October 31st, 2012 from JSTOR.

Lee, Donghoon. (February, 2013). *Household debt and credit: Student debt*. New York City, NY: Federal Reserve Bank of New York.

McMahon, Walter W. (February, 2009). *Higher learning, greater good: The private and social benefits of higher education*. Baltimore, MD: John Hopkins University Press

Myers, Richard S. and Pyles, Michelle R. (November, 1992). Relationships among High School Grades, ACT Test Scores, and College Grades. Mid-South Educational Research Association. Retrieved July 1st, 2011, from ProQuest database

National Center for Education Statistics. (May, 2011). *The condition of education 2011* (NCES 2011-03). Washington, D.C.: U.S. Department of Education

The National Center for Public Policy and Higher Education and Public Agenda. (February, 2010). *Squeeze Play 2010 Continued public anxiety on cost, harsher judgments on how colleges are run*. Washington, D.C.

Noftle, Erik E., and Robins, Richard W. (July, 2007). Personality predictors of academic outcomes: big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, Vol 93(1), Jul, 2007. pp. 116-130. Retrieved April 8th, 2013 from

EBSCO host database.

Pedrini, Bonnie C., and Pedrini, D. T. (January, 1974). The usefulness in predicting achievement and attrition among disadvantaged and regular freshman: a survey and study. University of Nebraska – Lincoln. Retrieved July 1st, 2011, from ProQuest database

Pew Charitable Trusts. (January, 2013). *How much protection does a college degree afford? The impact of the recession on recent college graduates*. Washington, D.C.

Raum, Tom. (April 2012). Student debt could be in crisis; \$1 trillion outstanding as college costs increase and jobs are harder to come by. St. Louis Post – Dispatch. Retrieved October 31st, 2012, from ProQuest database.

Siegfried, John J., Sanderson, Allen R., and McHenry, Peter. (March 2008). The economic impact of colleges and universities. *Change*, volume 40 issue 2, p. 24-31. Retrieved October 31st, 2012, from Ebsco host.

Stevens, Point. (August, 1969). Relationship of Academic Predictors to College Success. Wisconsin State University. Retrieved July 1st, 2011, from ProQuest database

Synco, Tracee M. (April 2012). Background or experience? Using logistic regression to predict college retention. The University of Alabama at Birmingham. Retrieved

October 31st, 2011 from ProQuest database.

TABLE 1. Descriptive Statistics of Variables

Dependent Variable	M	F	T	p-value	M	F	T
	Mean				Standard Deviation		
COLLEGE GPA	2.4	2.69	2.56	0.00195	1.2	1.12	1.17
	Frequency				Percentage		
RETURN	171	216	387	0.08544	58.56%	65.26%	62.12%
COLLEGE GPA > 2.0*	203	261	464	0.00758	69.52%	78.85%	74.00%
Independent Variables	M	F	T	p-value	M	F	T
	Mean				Standard Deviation		
ACT	21.33	21.27	21.3	0.86245	4.08	4.09	4.08
AGE	18.67	18.51	18.59	0.29561	2.09	1.88	1.98
HS GPA	3.18	3.36	3.27	<0.00001	0.5315	0.5124	0.5295
HS RANK	58.67	66.98	63.09	<0.00001	23.05	21.67	22.71
	Frequency				Percentage		
MAJOR	233	261	494	0.77182	79.79%	78.85%	79.29%

The inclusion of the p-value is to indicate which variables have a statistical difference between the male and female observations.

* Though placed under the dependent variable category due to it being a specified value of COLLEGE GPA, COLLEGE GPA > 2.0 was not used as a dependent variable in any of the regression analyses. It is here for purely descriptive purposes only.

Table 2. COLLEGE GPA Regression Analysis

Dependent Variable		COLLEGE GPA		
Multiple R				0.614945177
R Square				0.378157571
Adjusted R Square				0.372100664
Standard Error				0.925396734
Observations				623
F Score				62.43410787
Significance F				1.96701E-60
	Coefficients	Standard Error	t Stat	P-value
Intercept	-2.952599882	0.568772505	-5.19118	2.842E-07
AGE	0.060166609	0.019217937	3.130753	0.00182614
ACT SCORE	0.004442412	0.010832158	0.410113	0.681865468
HS RANK	0.001841722	0.004171717	0.441478	0.659021793
HS GPA	1.284126458	0.182033842	7.054328	4.6792E-12
GENDER	0.046895305	0.076181979	0.61557	0.538406028
MAJOR STATUS	-0.063123726	0.091677526	-0.68854	0.49137145

Bolded figures significant at $\alpha = .01$

Table 3. COLLEGE GPA Regression Analysis with AGE squared

Multiple R	0.618661778
R Square	0.382742395
Adjusted R Square	0.375716699
Standard Error	0.922728234
Observations	623
F Test	54.47750344
Significance F	1.7202E-60

	Coefficients	Standard Error	t Stat	P-value
Intercept	-6.601671854	1.799056762	-3.66951838	0.000264098
AGE	0.364126743	0.143502003	2.537433169	<u>0.011412703</u>
AGE^2	-0.005795925	0.002711796	-2.137300932	<u>0.032967284</u>
ACT COMP	0.004034032	0.010802612	0.373431144	0.70895627
HS RANK	0.002246618	0.004163999	0.539533716	0.589713919
HS GPA	1.281575216	0.18151285	7.060520615	<u>4.49718E-12</u>
GENDER	0.054098457	0.076037025	0.711475195	0.477059586
MAJOR	-0.047838235	0.091692497	-0.521724641	0.602049903

Bolded and underlined figures significant at $\alpha = .01$

Bolded figures significant at $\alpha = .05$

Table 4. RETURN Logistic Regression Analysis

Parameter	Estimate	Standard Error	P > ChiSq
Intercept	-4.478	1.3617	0.001
AGE	0.0725	0.0461	0.1153
ACT SCORE	-0.00583	0.0259	0.8221
HS RANK	0.0165	0.00962	0.0871
HS GPA	0.08559	0.4234	<u>0.0432</u>
GENDER	0.0198	0.1811	0.9129
MAJOR	-0.0526	0.2176	0.8092

Bolded p-values significant at $\alpha = .10$

Bolded and underlined p-values significant at $\alpha = .05$

Table 5. RETURN Logistic Regression Analysis with AGE squared

Parameter	Estimate	Standard Error	P > ChiSq
Intercept	-6.7223	4.1978	0.1093
AGE	0.2589	0.3325	0.4362
AGE Sqr	-0.00354	0.00624	0.5705
ACT SCORE	-0.00617	0.0259	0.8119
HS RANK	0.0167	0.00963	0.0822
HS GPA	0.8553	0.4234	<u>0.0434</u>
GENDER	0.0254	0.1814	0.8885
MAJOR	-0.0442	0.2182	0.8396

Bolded p-values significant at $\alpha = .10$

Bolded and underlined p-values significant at $\alpha = .05$

Table 6. Descriptive Statistics of the Admissions Decision Criteria

	M	F	T	p-value	M	F	T
	Frequency				Percentage		
ACT >= 21	166	182	348	0.4837	56.85%	54.98%	55.86%
HS RANK >= 50%	182	250	432	0.00036	62.33%	75.53%	69.34%
HS GPA >= 2.25	281	322	603	0.4593	96.23%	97.28%	96.79%
ACT + HS RANK	134	158	292	0.64552	45.89%	47.73%	46.87%
ACT + HS GPA	182	250	432	0.00036	62.33%	75.53%	69.34%
HS RANK + HS GPA	163	179	342	0.65994	55.82%	54.08%	54.90%
ADC							
Meets 0	8	6	14	0.4354	2.74%	1.81%	2.25%
Meets 1	73	54	127	0.00714	25.00%	16.31%	20.39%
Meets 2	77	113	190	0.03572	26.37%	34.14%	30.50%
Meets 3	134	158	292	0.64552	45.89%	47.73%	46.87%

The inclusion of the p-value is to indicate which variables have a statistical difference between the male and female observations.

Bolded and italicized figures significant at a = .01

Table 7. RETURN Logistic Regression Analysis with ADC

Parameter	Estimate	Standard Error	P > ChiSq
Intercept	-2.1055	0.9195	0.022
AGE	0.052	0.0446	0.2438
GENDER	0.221	0.1726	0.2003
MAJOR	0.0132	0.2127	0.9504
ADC	0.6934	0.1047	< 0.0001

Bolded p-values significant at $\alpha = .01$

Table 8. RETURN Logistic Regression Analysis with ADC and AGE squared

Parameter	Estimate	Standard Error	P > ChiSq
Intercept	-2.6336	4.0214	0.5125
AGE	0.0958	0.3281	0.7703
AGE Sqr	-0.00083	0.00618	0.8926
GENDER	0.2226	0.173	0.1981
MAJOR	0.0154	0.2133	0.9426
ADC	0.6942	0.1049	< 0.0001

Bolded p-values significant at $\alpha = .01$