VALUE OF SAT SCORES IN PREDICTING COLLEGE GPA

Sema Barlas

Executive Summary

Twenty first century knowledge economy demands every individual to have at least a college degree. Therefore, more students have been seeking admissions at universities. The problem with providing seats to a larger portion of student population is, however, that fewer students who are admitted to universities end up graduating in comparison to historical standards. By using data obtained from 105 computer science graduates, this study evaluated the value of commonly implemented admission policy of selecting students on the basis of Scholastic Aptitude Test (SAT) scores. Results suggest that the ability of SAT scores to predict college success is limited, especially for programs that attract students with relatively high SAT scores.

Background

Post-secondary degree has become more valuable than ever (Acemoglu and Author, 2010). Full-time employment rates for adults between the ages 25–34 in 2010 were 74% and 41% for individuals with and without post-secondary degree, respectively[1]. In 2010, the median individual with a degree from a higher education institution earned $27,000 more than the median individual with only a high school degree. Comparatively, in 1996 this difference was $16,500. This economic advantage due to post-secondary degree has been fueling a steady increase in demand for higher education. Undergraduate enrollment in degree-granting post-

[1] Statistics are obtained from National Center for Education Statistics (2011), unless stated otherwise.

secondary institutions grew from 34% of high school graduates in 2000 to 70% in 2009, or 17.6 million students. This number is expected to grow to 19.6 million in 2020.

However, studies also show increasing un-productivity in higher education institutions in that, in comparison to historical standards, both college graduation rates and improvement in even the basic skills of college graduates have been declining, and at the same time, higher education expenditures have been rising to become 2.6% of GDP in 2013 (Harris, 2013). National Center for Education Statistics reported 59% graduation rate in 2011 among full-time, first-time undergraduate students who began their pursuit of a bachelor's degree at a 4-year degree-granting institution in fall 2005. That is, large numbers of students occupy seats at the higher education institutions and incur the cost of college education but never graduate. Therefore, there is a growing interest in measurements predicting college success.

The Scholastic Aptitude Test (SAT) was developed by Carl Brigham in 1926 as a tool to allow universities to select students who are most likely to succeed. SAT scores have been used extensively since its inception in student admissions by universities and have been criticized for its bias against minority groups and women over the years. However, the value of SAT scores in predicting college success above and beyond high school GPA has not been questioned extensively. Given that SAT requires additional expense and effort, it is important to evaluate incremental value of SAT scores in predicting college success. Thus, the purpose of this project is to evaluate incremental benefit of SAT scores in predicting college success after controlling for high school GPA.

College success is the dependent variable in this project and it is defined as the average GPA obtained at the university. The predictor variables are the high school GPA and scores from the mathematics and verbal sections of SAT. The SAT scores are believed to reflect general mathematics and verbal skills, whereas high school GPA is generally taken as a measure of students' understanding of actual content taught at the high school. In a way, high school GPA and SAT scores measure different skills and it is therefore possible that general skills are significant predictors of college success above and beyond the quality of high school content knowledge captured by high school GPA. Yet, it is also possible that the same factors that influence students' high school GPA also determine their SAT scores, making these two measurements highly redundant. This project thus tests the following hypotheses in order to evaluate the incremental value of SAT scores in predicting college success empirically.

H_0: $\beta_{math\_SAT} = 0 \; and \; \beta_{verbal\_SAT} = 0, after \; controlling \; for \; high \; school \; GPA.$

H_1: $\beta_{math\_SAT} \neq 0 \; or \; \beta_{verbal\_SAT} \neq 0, after \; controlling \; for \; high \; school \; GPA.$

Data

Data were collected as part of a research project conducted by Thomas W. MacFarland and reported in a case study prepared by Emily Zitek. The case study can be found in an online statistics book developed by Rice University, University of Houston Clear Lake, and Tufts University at http://onlinestatbook.com/.

The participants were 105 students who graduated from a state university with a B.S. in computer science. Although it is stated as such, it is puzzling that measurements were taken only from those who graduated. Provided that this was not a reporting mistake, the results of this project should be interpreted with caution since the distributions of the measurements are truncated.

The measurements are high school GPA (high_GPA), mathematics SAT score (math_SAT), verbal SAT score (verb_SAT), computer science GPA (comp_GPA) and overall GPA attained at the university (univ_GPA).

Results

Academic Profile of the Sample

College Board reported that in 2013 the average scores for mathematics and verbal sections of SAT were 514 and 469, respectively. Table 1 provides the central tendency and standard deviation for all measures obtained from the sample. Average mathematics and verbal SAT scores of sample are much higher than of those students who took SAT in 2013. According to the one sample t test, sample mean mathematics, $t(104) = 20.79$, $p<.001$, and verbal SAT

scores, ($t(104) = 16.70$, $p<.001$), are significantly different from the corresponding population means. Figure 1 presents the distribution of mathematics and verbal SAT scores, as well as 95% upper confidence interval for means. That is, the sample consists of higher achieving students relative to population of students who took SAT in 2013.

Table 1

Descriptive Statistics

	Mean	STD	Median	Mode
High_GPA	3.07	0.52	3.17	3.28
Math_SAT	623	53.76	612	554
Verb_SAT	598	62.96	591	538
Comp_GPA	3.13	0.51	3.21	3.17
Univ_GPA	3.17	0.45	3.29	3.28

Figure 1

Distribution of Mathematics and Verbal SAT Scores

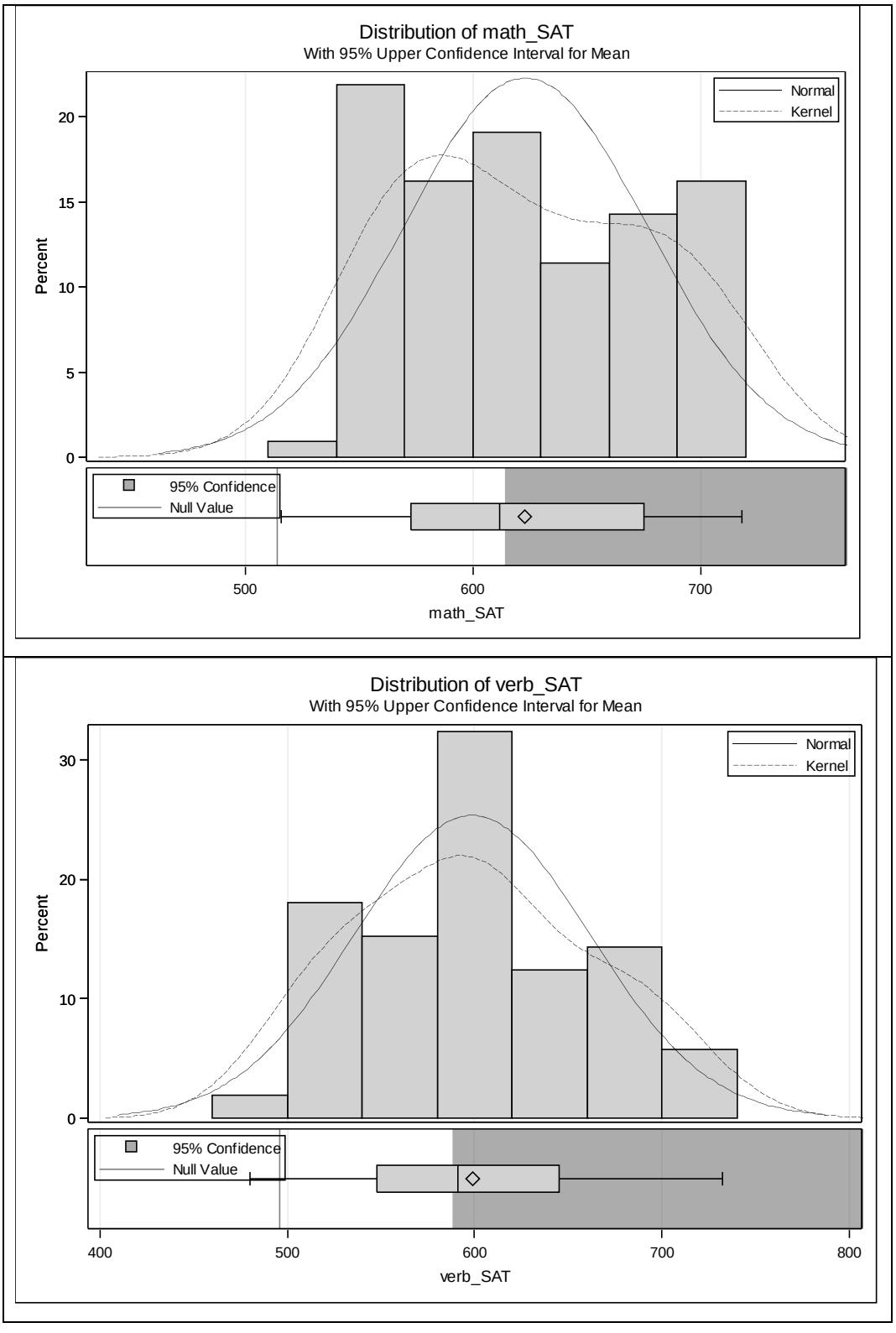

Distribution of math_SAT
With 95% Upper Confidence Interval for Mean
Normal
Kernel
Percent
20
15
10
5
0
95% Confidence
Null Value
500
600
700
math_SAT
Distribution of verb_SAT
With 95% Upper Confidence Interval for Mean
Normal
Kernel
Percent
30
20
10
0
95% Confidence
Null Value
400
500
600
700
800
verb_SAT

Table 2 presents the correlations of the predictor variables with the university and computer science GPAs. The dependent variable, university GPA, has the highest correlation (0.78) with the high school GPA and moderate correlations with the mathematics (0.66) and verbal (0.65) SAT scores. More importantly, the high school GPA correlates with mathematics (0.77) and verbal (0.73) SAT scores highly, limiting the ability of SAT scores in predicting university GPA.

Table 2

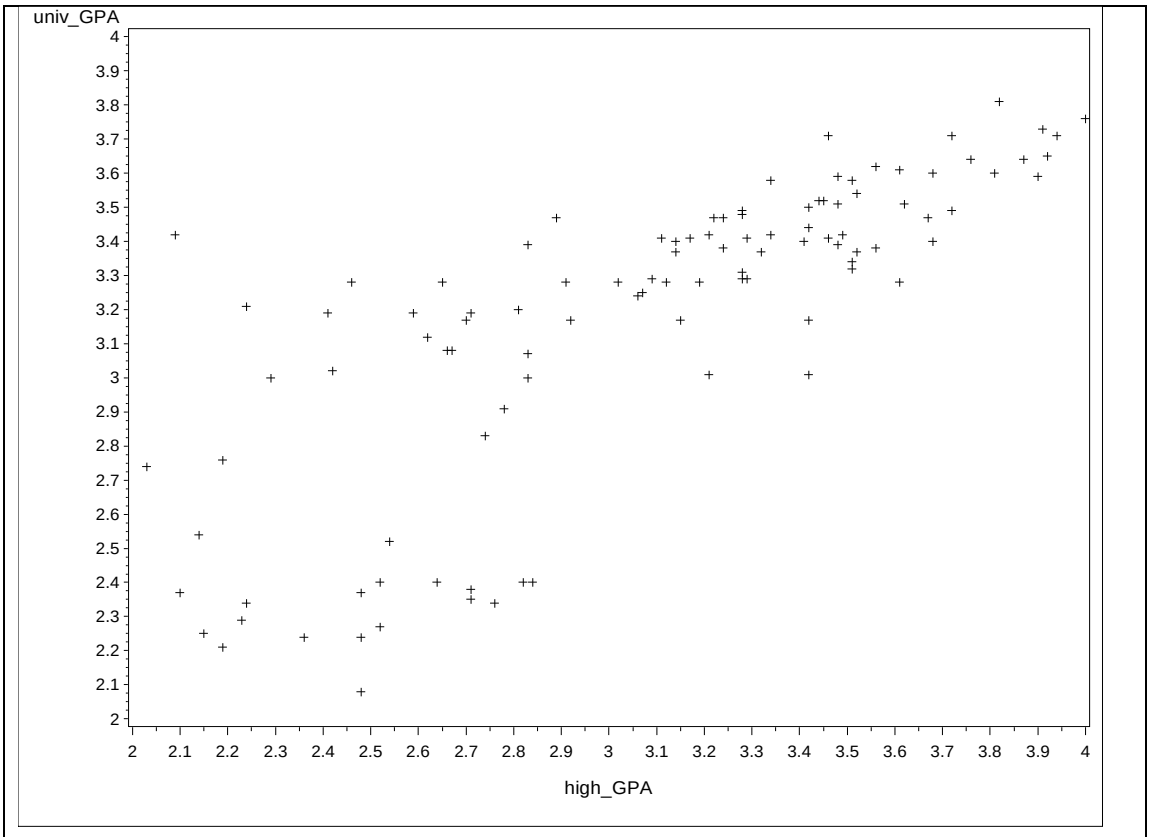
Correlations between Target and Predictor Variables

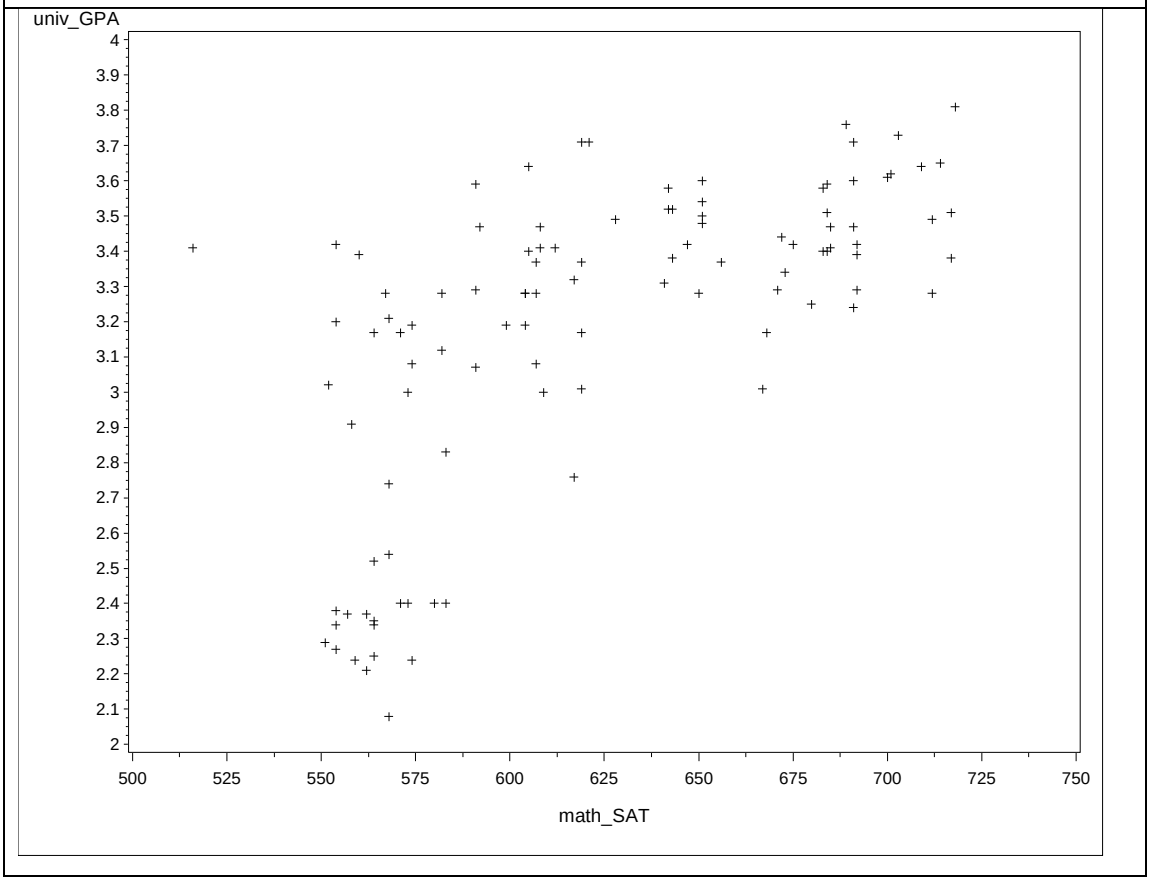
	High_GPA	Math_SAT	Verb_SAT	Comp_GPA
Comp_GPA	0.79	0.69	0.64	1.00
Univ_GPA	0.78	0.66	0.65	0.93

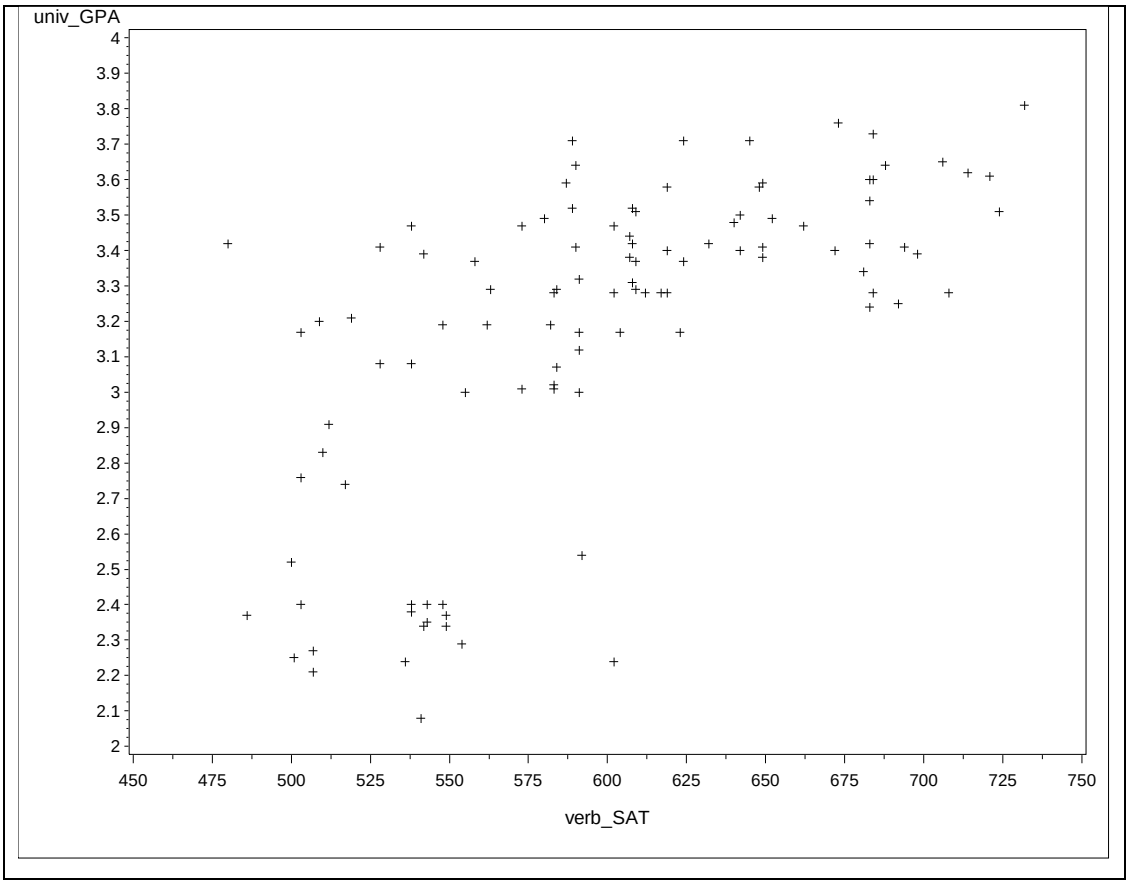
Figure 2 displays the scatter plots of university GPA with all three predictor variables. Whereas the relation between university GPA and high school GPA appears linear, there seems to be greater variance in university GPAs of students with low high school GPA. This may result in the problem of heteroscedasticity in regression analysis. Also, the relationship between university GPA and mathematics and verbal SAT scores seems slightly curvilinear; therefore, quadratic terms for both SAT scores will be included in the regression model for predicting university GPA.

Figure 2

The Relationship between University GPA and Predictor Variables.


univ_GPA
4
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
3
2.9
2.8
2.7
2.6
2.5
2.4
2.3
2.2
2.1
2
2 2.1 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 3 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 4
high_GPA


univ_GPA
4
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
3
2.9
2.8
2.7
2.6
2.5
2.4
2.3
2.2
2.1
2
500 525 550 575 600 625 650 675 700 725 750
math_SAT



Value of SAT Scores in Predicting University Success

To assess incremental value of SAT scores in predicting college success, a regression model is developed where the university GPA is the dependent variable and the high school GPA and SAT scores are the predictors. Squared mathematics and verbal SAT scores are also included in the model as predictors to accommodate the slightly nonlinear relationship between university GPA and SAT scores. Stepwise regression procedure is utilized in variable selection where only those variables that exceed 0.15 significance level are allowed to remain in the model. The final model included high school GPA and verbal SAT as predictor variables with 62% of explained variance in university GPA. However, diagnostic analysis identified an outlier with Cook's Distance measure of 0.216. Elimination of this observation resulted in a model with the same predictor variables and 67% of explained variance in university GPA. Table 3 presents the parameter estimates resulting from this analysis. The table also includes heteroscedasticity

consistent standard errors, which will be used in testing the hypotheses, because of the closing fan structure of the residuals indicating the violation of homogeneity of variances assumption.

Results suggest that high school GPA is the best predictor of university GPA, $t(1) = 8.61$, $p<.0001$, accounting for 65% of the variance in university GPA. As expected, the higher the high school GPA is, the higher the university GPA will be. Furthermore, only the verbal SAT score makes into the model, $t(1) = 2.92$, $p<0.004$, with 2% incremental explained variance in university GPA. That is, although it is small in magnitude, the ability of verbal SAT scores to predict college success is significant.

Table 3

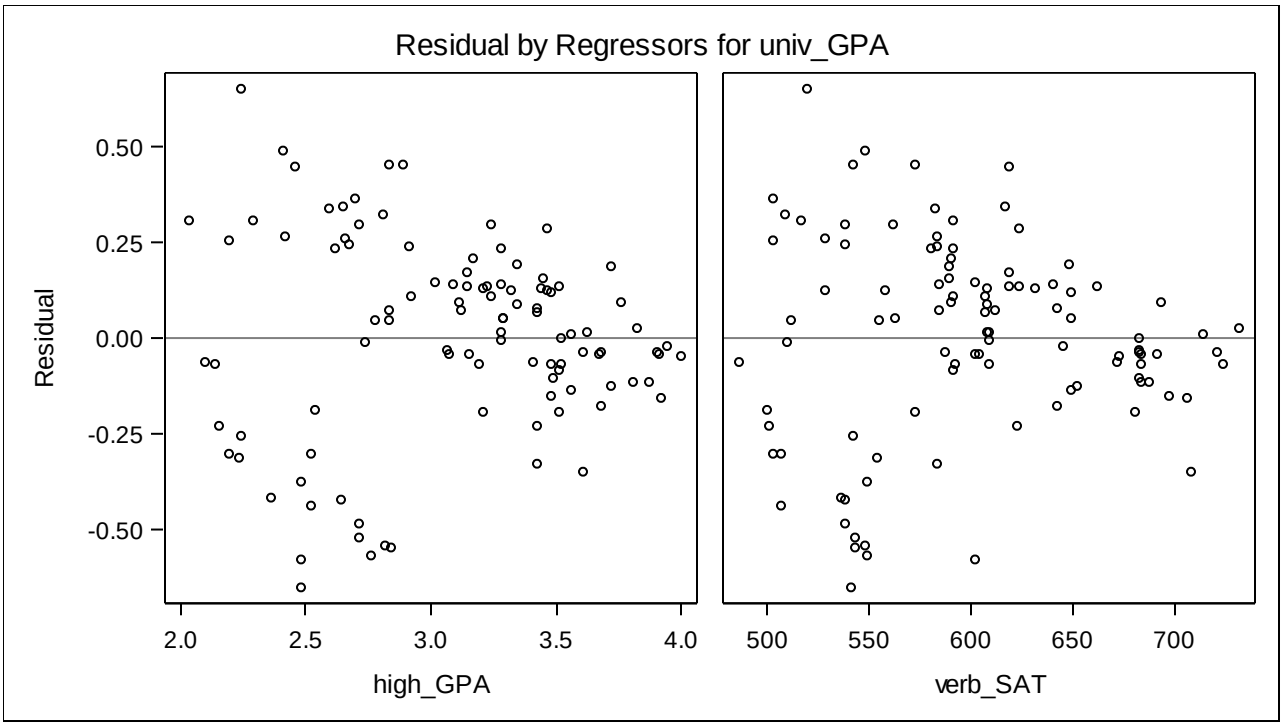
Predicting University Success on the Basis of SAT Scores After Controlling for High School GPA.

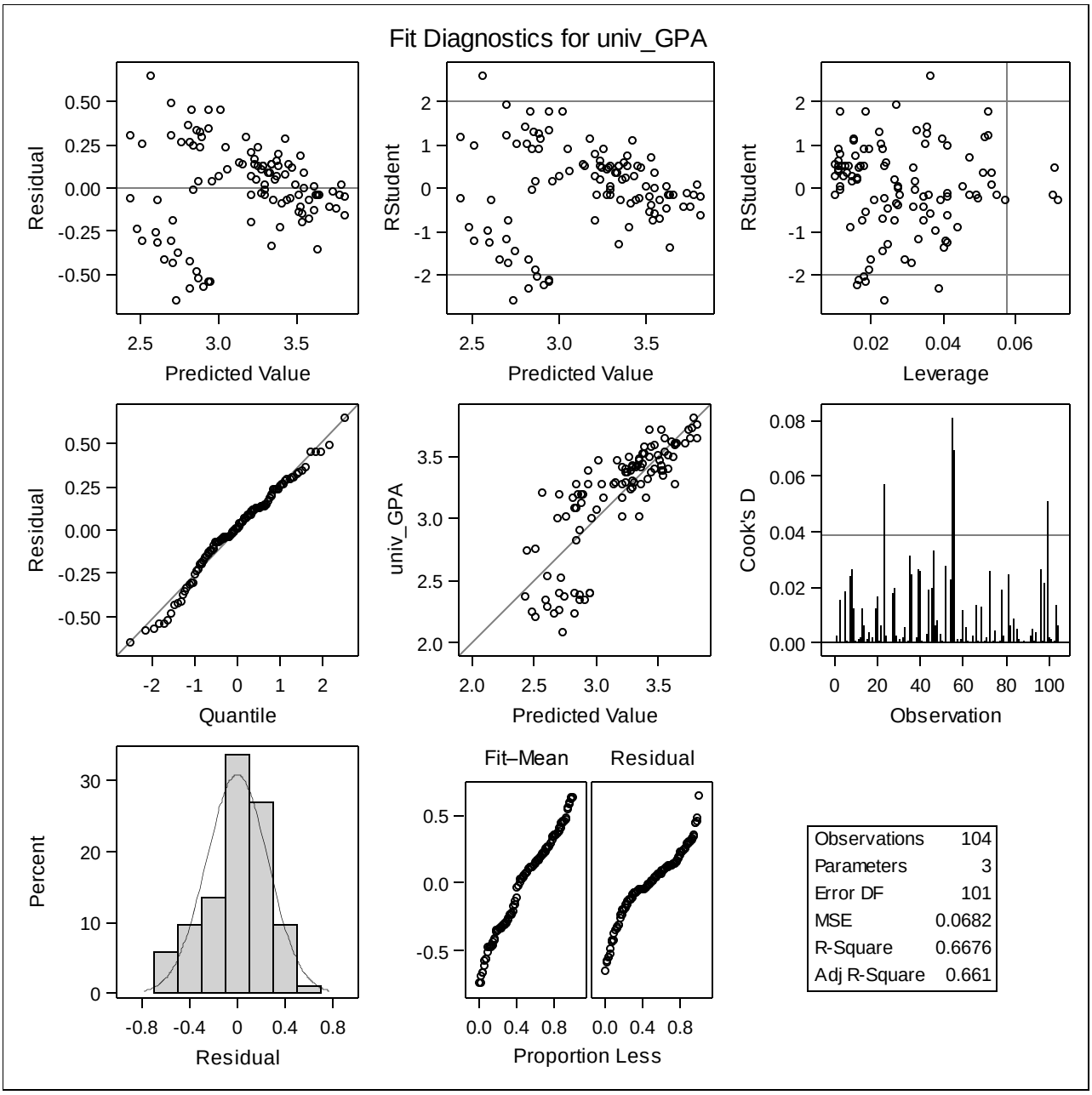
Parameter Estimates									
						Heteroscedasticity Consistent			
Variable	*DF*	*Parameter Estimate*	*Standard Error*	*t Value*	*Pr > |t|*	*Standard Error*	*t Value*	*Pr > |t|*	*Standardized Estimate*
Intercept	1	0.51225	0.25162	2.04	0.0444	0.24643	2.08	0.0402	0
high_GPA	1	0.58420	0.07235	8.08	<.0001	0.06788	8.61	<.0001	0.66381
verb_SAT	1	0.00143	0.00059328	2.40	0.0180	0.00048847	2.92	0.0043	0.19764

Diagnostic analysis presented in Figure 3 confirms the problem of heteroscedasticity. Yet, linearity assumption and assumptions with respect to the error distribution seem to hold.

Figure 3

Regression Diagnostics


Residual by Regressors for univ_GPA
Residual
0.50
0.25
0.00
-0.25
-0.50
2.0
2.5
3.0
3.5
4.0
high_GPA
500
550
600
650
700
verb_SAT



Discussion

Overall, results suggest that, although it is significant, the value of SAT scores in predicting college performance after controlling for high school GPA is very small. Only verbal SAT scores improve predictions significantly, explaining about 2% of incremental variance in university GPA. Given the amount of effort and cost involved in the administration of SAT, universities may need to question the benefits of demanding SAT scores from students and using

these scores in admission decisions. Yet, 2% improvement in predicting student success may be considered significant especially by large universities.

However, the results of this project should be interpreted with caution since it is not clear whether the data included all students or only the students who graduated. Focusing on only the graduating students will restrict the range and result in underestimation of the value of SAT scores in predicting university success. Also, the sample was limited to computer science major that may attract relatively more high achieving students in comparison to the population of applicants. Indeed, statistical analysis confirmed that the participants have significantly higher mathematics and verbal SAT scores in comparison to the population of SAT takers. It is possible that SAT scores do not differentiate well among high achievers but they may be most useful in identifying high risk students who may not graduate among relatively low achieving applicants.

In conclusion, the ability of SAT scores to predict college success appears to be limited, especially for programs that attract students with relatively high SAT scores. Given the limitations of the data and unrepresentative nature of the sample, however, the findings of this study should be regarded as tentative. Perhaps the most useful conclusion suggested by the results is that universities need to investigate the benefits of their admission policies systematically and improve their policies on the basis of real empirical evidence.

References

Acemoglu, D. and Author, D. H. (2010). Skills, tasks, and technologies. Implications for employment and earnings. In Orley Ashenfelter and David Cards, Eds., *Handbook of Labor Economics*, Volume **4**, Amsterdam: Elsevier: North Holland.

Harris, N. D. (2013). Addressing the declining productivity of higher education using cost effectiveness analysis. American Enterprise Institute.