

ASSESSMENT CORRELATIONS

Prepared for Washtenaw Intermediate
School District

May 2015



In the following document, Hanover Research uses data provided by Washtenaw Intermediate School District to determine the correlations between ACT, PLAN, PSAT, and SAT assessments.



www.hanoverresearch.com

TABLE OF CONTENTS

Executive Summary and Key Findings	3
INTRODUCTION	3
KEY FINDINGS.....	3
Section I: Data and Methodology.....	5
DATA.....	5
METHODOLOGY	6
Correlation Analysis	6
Linear Regression Analysis.....	7
Section II: Correlation Analysis	9
SUMMARY OF FINDINGS	9
PRELIMINARY ASSESSMENTS	9
CORRELATIONS BETWEEN PRELIMINARY ASSESSMENTS AND THE ACT.....	10
SUMMATIVE ASSESSMENTS	10
Section III: Regression Analysis	12
SUMMARY OF FINDINGS	12
PSAT AND PLAN ASSESSMENTS	12
PRELIMINARY (PSAT AND PLAN) ASSESSMENTS AND THE ACT.....	13
SAT AND ACT ASSESSMENTS	15
Appendix: Correspondence between Actual and Predicted Scores.....	17

EXECUTIVE SUMMARY AND KEY FINDINGS

INTRODUCTION

In this report, Hanover Research evaluates the extent of correlations between different assessment outcomes at Washtenaw Intermediate School District (WISD). We use correlation matrices and linear regression analysis to find the extent of correlation between PLAN and PSAT scores, PLAN and ACT scores, PSAT and ACT scores, and SAT and ACT scores.

We find that all the assessment outcomes are highly correlated, with the strongest correlations between the ACT and SAT assessments, and between the PSAT and ACT assessments.

This report consists of three sections:

- **Section I** outlines the data provided by WISD, the data processing conducted by Hanover Research, and the methodologies employed in the analyses.
- **Section II** presents a correlation analysis of the composite and subject assessments.
- **Section III** provides the results of linear regression analysis of the relationship between assessment composite scores.

KEY FINDINGS

- **All the assessment outcomes we compare are highly correlated with each other.**
 - With models that only include scores from a single assessment and the squared term of that assessment, the models explain between 64 and 84 percent of the variation in the predicted assessment.
 - The results of the correlation matrices and the linear regression analysis all show similarly strong correlations between the analyzed assessments.
- **Students who score highly on either the PSAT *or* PLAN are expected to also score highly on the ACT assessment.** Both the PSAT and PLAN assessments are highly correlated with ACT performance; we find that the PSAT explains about 82 percent of ACT scores, while the PLAN explains about 75 percent of ACT scores. In other words, perhaps somewhat surprisingly, the difference in model R-squared suggests that the PSAT is actually more reliable than the PLAN assessment in predicting successful students on the ACT assessment. This should not be taken as generalizable more broadly, beyond the dataset supplied to Hanover by Washtenaw ISD. *The Appendix contains a chart showing detailed correspondence between actual and predicted scores.*
- **ACT performance is highly correlated (at nearly 84 percent) with SAT performance.** Because of the high level of correlation between the PSAT/PLAN and ACT, we can translate these correlations to SAT performance. This means that students who score highly on the PSAT or PLAN are also expected to score highly on the SAT performance.

However, due to the very small sample size of students who have taken any combination of SAT and other assessments, we have less confidence in our SAT estimates than our ACT estimates.

- **Overall, the strongest correlations are between the PSAT and ACT composite scores** (correlation coefficient = 0.895) **and between the SAT and ACT composite scores** (correlation coefficient = 0.849).
- The **strongest subject score correlations** are between the **PSAT writing and the ACT English** (correlation coefficient = 0.843), between the **SAT math and ACT science** scores (correlation coefficient = 0.827), and between the **PSAT math and the ACT math** (correlation coefficient = 0.822).
- Models that include nonlinear transformations of the predicting variable fit the data better than models that only describe the linear relationship between the predicting and predicted variables.

SECTION I: DATA AND METHODOLOGY

In this section, Hanover Research discusses the data analyzed in this report and presents our methodological approaches to the analysis.

DATA

Washtenaw Intermediate School District (WISD) provided Hanover four types of student-level test score data for two high schools, Chelsea High School (CHS) and Saline High Schools (SHS), covering the following assessments: ACT, PSAT, PLAN, and SAT. It is important to note that not all students sat for at least two assessments. As a result, only students with at least two assessments are included in the final analysis. The following lists the number of such students by assessment type.

ACT Assessment Data include subject (English, mathematics, reading, and science) and composite test scores in 2012-13 and 2013-14 for 792 students, 378 in CHS and 414 in SHS.

PSAT Assessment Data include test scores in critical reasoning, mathematics, and writing skills for 527 students in CHS and 326 students in SHS between 2010-11 and 2013-14. However, PSAT data for Saline High School are not available for 2011-12.

PLAN Assessment Data include subject (English, mathematics, reading, and science) and composite test scores for 666 students in CHS between 2010-11 and 2013-14, and 563 students in SHS in 2012-13 and 2013-14.

SAT Assessment Data are available only for 40 students in CHS. Note that SAT scores are available for CHS students who are expected to graduate in 2012-13, whereas SAT data available for SHS students who took the assessment between May of 2009 and November of 2014. As such, these assessments could not be linked to available PSAT, PLAN, or ACT assessments.

Hanover combines these data into a single analytic file which represents student-level data for 1,246 students who have outcomes for at least two assessments. These assessments are normed by year and transformed into percentile scores for the later analyses. Figure 1.1 shows the means, standard deviations, and counts of scaled assessment outcomes for students who have at least two assessments to evaluate.

Figure 1.1: Summaries of Assessment Scores

ASSESSMENT	CHELSEA HIGH SCHOOL			SALINE HIGH SCHOOL		
	Mean	SD	Count	Mean	SD	Count
ACT English	22	6	378	24	6	414
ACT Math	24	5	378	24	6	414
ACT Reading	23	6	378	24	6	414
ACT Science	23	5	378	24	5	414
ACT Writing	21	5	377	23	6	414
ACT Composite	23	5	378	24	5	414
PLAN English	20	4	666	19	5	558
PLAN Math	22	5	666	21	6	560
PLAN Reading	20	5	666	20	5	563
PLAN Science	21	4	666	20	5	563
PLAN Composite	21	4	666	20	5	563
SAT Reading	625	86	40	-	-	-
SAT Math	612	90	40	-	-	-
SAT Writing	598	89	40	-	-	-
SAT Composite	1835	234	40	-	-	-
SAT Verbal and Math	1237	159	40	-	-	-
PSAT Critical Reasoning	50	9	527	56	10	326
PSAT Math	53	9	527	59	9	326
PSAT Writing	48	10	527	54	10	326
PSAT Composite	151	26	527	170	26	326

METHODOLOGY

Hanover constructs correlation matrices to describe the statistical relationship between pairwise combinations of composite and subject scores. In addition, we conduct regression analysis to estimate the relationship between composite scores while allowing the assessments to follow a non-linear relationship. Specifically, we estimate the quadratic relationship between ACT and PSAT assessment scores, ACT and PLAN assessment scores, PLAN and PSAT assessment scores, and SAT and ACT assessment scores.

It is important to note that any correlations that are found are not necessarily generalizable to *other* schools or other times. This report only explores potential empirical relationships in WISD’s data, and does not purport to examine the relationship between these assessments more broadly.

CORRELATION ANALYSIS

In Section II, Hanover presents correlation matrices for the assessment outcomes, including the assessment subject scores. These correlations can range from -1 to 1, with -1

representing perfect negative correlation and 1 representing perfect positive correlation. Zero represents no correlation between the variables. **Significance “asterisks” represent the level of statistical significance of any correlation**, based on a simple t-test of whether the correlation is zero.

LINEAR REGRESSION ANALYSIS

To further evaluate the relationship between the assessment composite outcomes, Section III uses linear (ordinary least squares) regression models. The reason for using this type of model is that it facilitates meaningful comparisons between groups by allowing the inclusion of additional variables. Regression models allow us to include additional data, including control variables and transformations of the predicting variables. The primary reason for including regression analysis in the present project is that regression analysis allows Hanover to investigate potential non-linear relationships between variables by including the square of the predicting assessment outcome. For ease of interpretation, we transform all assessment test scores into percentile ranks for each student at the district. This allows us to compare between assessments that are not measured on the same scale and provide a uniform scale for the analysis.

The formal representation of the regression model is such that each model has a single outcome variable and a set of predictor variables which include the predicting variables and the square of the predicting variable. For each outcome variable, we estimate the following regression equation:

$$Y_i = \beta_0 + \beta_1 * (\text{PredictingScore}_i) + \beta_2 * (\text{PredictingScore}_i)^2 + \epsilon_i \quad [1]$$

Y_{it} denotes the outcome variable which is an assessment score for student i . Predicting Score is the given predicting variable, and ϵ_i is the idiosyncratic error term. For instance, in a regression where the outcome variable is the ACT test score and the predicting variable is the PLAN test score, the regression equation would be represented as follows:

$$\text{ACT}_i = \beta_0 + \beta_1 * \text{PLAN}_i + \beta_2 * \text{PLAN}_i^2 + \epsilon_i \quad [2]$$

The parameters of interest to the evaluation are β_1 and β_2 , which estimate the quadratic relationship between the outcome assessment and the predicting assessment. Further, the resulting coefficients enable us to compute the expected outcome score given a certain score on the predicting assessment. A positive and statistically significant estimate of β_1 indicates that there is a positive correlation between the scores, and a positive and statistically significant estimate of β_2 indicates that the expected outcome score increases at an increasing rate when moving from lower values to higher values of the predicting variable. Thus, a negative value of β_2 indicates that the expected outcome score increases at a slower pace at higher values of the predicting assessment.

Finally, in order to reduce the likelihood that different student populations taking different assessments contributes to greater relative correlation between one set of assessments and

another, Hanover restricts the regression models in which PLAN and PSAT scores are used to predict ACT scores to students who have both PLAN and PSAT, as well as ACT scores. This enables us to compare the relative predictive power of the PSAT and the PLAN assessments in relation to the ACT assessment.

SECTION II: CORRELATION ANALYSIS

This section presents the results of Hanover’s analysis of correlations between assessment outcomes, including subject scores.

SUMMARY OF FINDINGS

- There are strong correlations between all the assessment outcomes we compare.
- The PSAT scores correlate slightly more strongly with the ACT scores than do the PLAN scores.
- The strongest correlations are between the PSAT and ACT composite scores (correlation coefficient = 0.895)¹ and between the SAT and ACT composite scores (0.849).
- Of the subject score correlations, the strongest are between the PSAT writing and the ACT English (0.843), between the SAT math and ACT science scores (0.827), and between the PSAT math and the ACT math (0.822).

PRELIMINARY ASSESSMENTS

The PSAT and PLAN assessments correlate with each other strongly, and even among the subject scores, the lowest correlation, which is between the PSAT math and the PLAN reading score, is 0.513 and is statistically significant at the 99 percent level. Figure 2.1 presents the correlations between PSAT and PLAN scores. The strongest correlations are between the PSAT composite and the PLAN composite (0.791), between the PSAT composite and the PLAN English (0.754), and between the PSAT writing and the PLAN English (0.740).

Figure 2.1: PSAT and PLAN Correlations

	PLAN ENGLISH	PLAN MATH	PLAN READING	PLAN SCIENCE	PLAN COMPOSITE
PSAT Critical Reasoning	0.699***	0.558***	0.684***	0.615***	0.722***
PSAT Math	0.589***	0.695***	0.513***	0.607***	0.687***
PSAT Writing	0.740***	0.553***	0.643***	0.596***	0.718***
PSAT Composite	0.754***	0.671***	0.683***	0.676***	0.791***
Number of Observations	831	833	836	836	836

Note: Asterisks denote statistical significance, as follows. *** p<0.01, ** p<0.05, * p<0.1

¹ All statistics in Section II refer to correlation coefficients.

CORRELATIONS BETWEEN PRELIMINARY ASSESSMENTS AND THE ACT

The preliminary assessments, the PLAN and the PSAT, correlate strongly with the later ACT assessment.

Figure 2.2 presents Pearson correlation coefficients between the ACT and the PLAN and PSAT. The PSAT composite correlates with the ACT composite to a somewhat higher degree (0.895) than the PLAN composite (0.865), but both strongly correlate with ACT composite scores. Among subject scores, the highest correlations are between PSAT writing and the ACT English (0.843), between the PSAT math and the ACT math (0.822), and between the PSAT critical reasoning and the ACT English (0.815).

It is notable that while there are high correlations between the PLAN subject tests and their ACT counterparts, a number of pairings between the PSAT and ACT subject tests have a higher correlation, despite the fact that the PLAN test is designed to simulate the ACT. For example, the correlation between PLAN science scores and ACT science scores is 0.724, while the correlation between the PSAT critical reasoning scores and the ACT science scores is 0.748.

Figure 2.2: PLAN and PSAT Correlations with ACT

	ACT ENGLISH	ACT MATH	ACT READING	ACT SCIENCE	ACT WRITING	ACT COMPOSITE
<i>PLAN Assessment</i>						
PLAN English	0.797***	0.643***	0.709***	0.680***	0.780***	0.783***
PLAN Math	0.688***	0.806***	0.603***	0.731***	0.680***	0.775***
PLAN Reading	0.700***	0.581***	0.748***	0.661***	0.699***	0.744***
PLAN Science	0.679***	0.698***	0.648***	0.724***	0.675***	0.754***
PLAN Composite	0.808***	0.776***	0.762***	0.790***	0.799***	0.865***
Number of Observations	775	775	775	775	774	775
<i>PSAT Assessment</i>						
PSAT Critical Reasoning	0.815***	0.671***	0.800***	0.748***	0.791***	0.845***
PSAT Math	0.711***	0.822***	0.638***	0.728***	0.698***	0.795***
PSAT Writing	0.843***	0.622***	0.761***	0.663***	0.812***	0.805***
PSAT Composite	0.868***	0.771***	0.805***	0.784***	0.845***	0.895***
Number of Observations	401	401	401	401	401	401

Note: Asterisks denote statistical significance, as follows. *** p<0.01, ** p<0.05, * p<0.1

SUMMATIVE ASSESSMENTS

Only 40 students have SAT scores and any other assessment, and of those, only 25 also have ACT scores. Despite the small sample size, it is clear that for the students who took both assessments, there is a strong correlation between the assessment scores. These results can be seen in Figure 2.3. The correlation coefficient for the two composites is 0.849, and it is statistically significant at the 99 percent level. The strongest subject score

correlations are, between the SAT math and ACT science scores (0.827), between the two assessments' math scores (0.815), and between the two assessments' writing scores (0.768).

Figure 2.3: SAT and ACT Correlations

	ACT ENGLISH	ACT MATH	ACT READING	ACT SCIENCE	ACT WRITING	ACT COMPOSITE
SAT Reading	0.663***	0.447*	0.733***	0.704***	0.663***	0.690***
SAT Math	0.743***	0.815***	0.720***	0.827***	0.707***	0.833***
SAT Writing	0.710***	0.587**	0.555**	0.580**	0.768***	0.633***
SAT Composite	0.772***	0.725***	0.793***	0.868***	0.727***	0.849***
Number of Observations	25	25	25	25	25	25

Note: Asterisks denote statistical significance, as follows. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

SECTION III: REGRESSION ANALYSIS

This section presents the results of Hanover's regression analysis of ACT and SAT outcome assessments.

SUMMARY OF FINDINGS

- **The composite (ACT and SAT) scores examined in this section are strongly predictive of each other.**²
 - With models that only include scores from a single assessment and the squared term of that assessment, the models explain between 64 and 84 percent of the variation in the predicted assessment.
- **The PSAT assessment scores are a better predictor of ACT scores** (R-squared = 0.819) **than the PLAN assessment scores** (R-squared = 0.756), using the same sample of students in each regression model.
- Models that include nonlinear transformations of the predicting variable fit the data better than models that only include the linear relationship between the predicting and predicted variables.³

PSAT AND PLAN ASSESSMENTS

The PSAT assessment and its squared term are strong predictors of the PLAN assessment score. The negative squared term, which can be seen in Figure 3.1, implies that PLAN scores increase at a decreasing rate as PSAT scores increase. This nonlinear quadratic relationship between PSAT and PLAN scores is depicted graphically in Figure 3.2. Additionally, the figure plots the 95 percent confidence interval around the predicted relationship between PLAN and PSAT.

The R-squared of the regression model represents the percent of the variation in PLAN scores which is explained by the PSAT scores. The R-squared for this model is 0.638, meaning that 63.8 percent of the variation in PLAN composite scores are explained by the PSAT composite score, and can be seen in Figure 3.1. The R-squared itself is telling of the correlation between the PSAT and PLAN assessments. However, to ascertain the level of correlation between the PSAT/PLAN assessments and students' academic aptitude, we estimate the relationship of each of PSAT and PLAN with ACT performance. Thus, we can compare both assessments in a controlled (identical) setting to examine whether the PSAT or the PLAN is better at preparing students for success on the ACT.

² SAT composite scores are simply the sum of critical reasoning, verbal, and math scores.

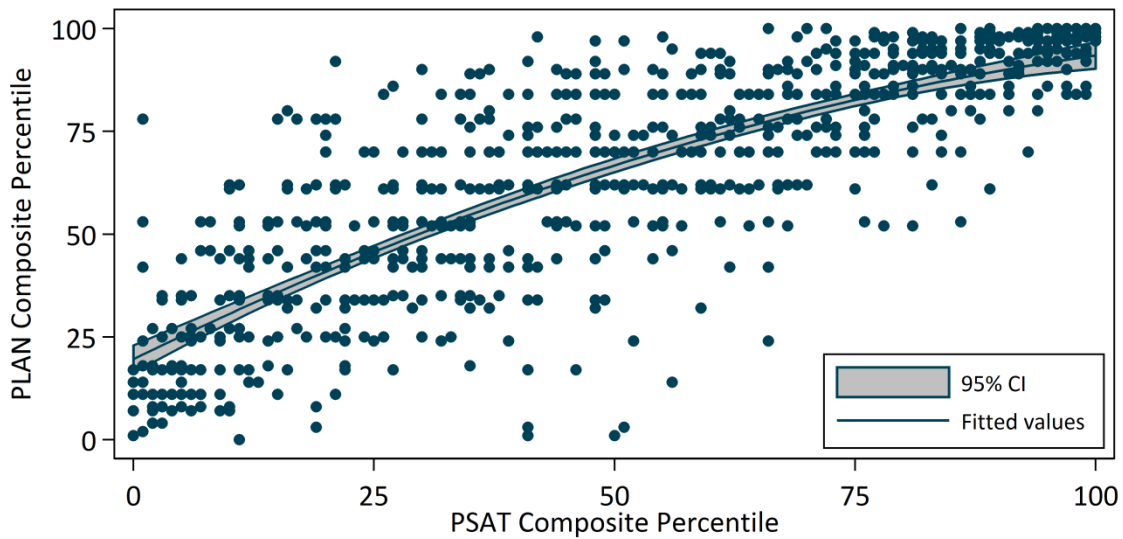
³ Model goodness of fit is determined using the model R-squared, which measures the percentage of the total variation in the dependent variable that is explained by the model.

Figure 3.1: PSAT Predicting PLAN, Regression Analysis

	COEFFICIENT	STANDARD ERROR
PSAT Composite Percentile	1.146***	(0.078)
PSAT Composite Percentile Squared	-0.004***	(0.001)
Constant	19.612***	(1.692)
Observations	836	
R-squared	0.638	

Notes: Coefficients are estimated using Ordinary Least Squares (OLS), using equation [1]. Numbers in parentheses denote standard errors. Asterisks denote statistical significance, as follows: *** p<0.01, ** p<0.05, * p<0.1.

Figure 3.2: PLAN and PSAT Scatterplot



PRELIMINARY (PSAT AND PLAN) ASSESSMENTS AND THE ACT

We find that The PLAN and PSAT assessment scores are strong predictors of ACT scores. The squared term of the PLAN test, which can be seen in Figure 3.3, is positive, indicating that ACT scores increase at an increasing rate as PLAN scores increase. This somewhat nonlinear relationship can be seen in Figure 3.4. For the PSAT, the squared term is negative, meaning that ACT scores increase at a decreasing rate as PSAT scores increase. This nonlinear relationship is depicted in Figure 3.5.

Figure 3.3: PSAT and PLAN Predicting ACT, Regression Analysis

VARIABLES	PLAN MODEL	PSAT MODEL
PLAN Composite Percentile	0.281**	
	(0.117)	
PLAN Composite Percentile Squared	0.005***	
	(0.001)	
PSAT Composite Percentile		1.279***
		(0.081)
PSAT Composite Percentile Squared		-0.005***
		(0.001)
Constant	17.663***	13.961***
	(3.130)	(1.698)
Observations	384	384
R-squared	0.756	0.819

Notes: Coefficients are estimated using Ordinary Least Squares (OLS), using equation [1]. Numbers in parentheses denote standard errors. Asterisks denote statistical significance, as follows: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The R-squared of the PLAN model is 0.756 and the R-squared of the PSAT model is 0.819. Since the only variables in the models are assessment scores and their squared terms, the greater R-squared from the PSAT model implies that the PSAT assessment is an overall better predictor of ACT scores than the PLAN assessment. This fits with the findings of Section II and is especially interesting since the PLAN assessment is designed to be similar to the ACT.

Figure 3.4: PLAN Predicting ACT Scatterplot

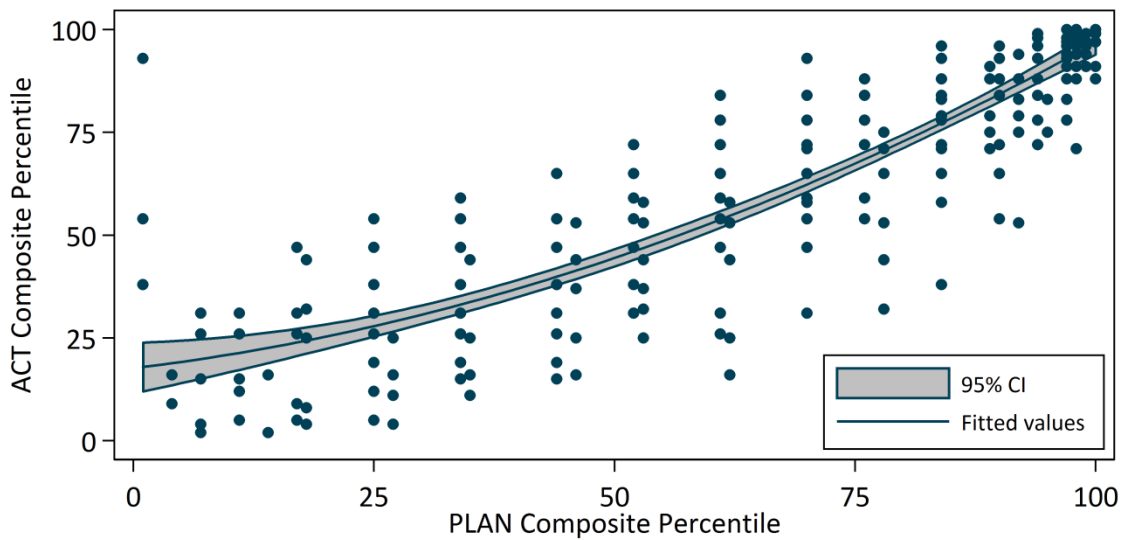
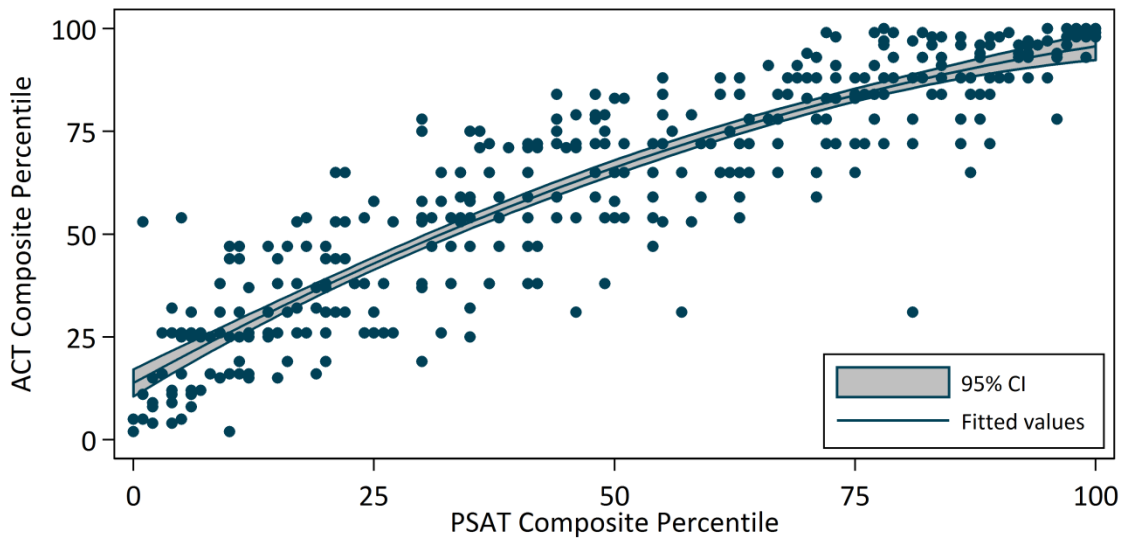


Figure 3.5: PSAT Predicting ACT Scatterplot



SAT AND ACT ASSESSMENTS

Even with small sample sizes, ACT performance is strongly predictive of SAT performance, and this relationship is statistically significant at the 99 percent level. Like the previous models discussed, the squared term, which can be seen in Figure 3.6, is also statistically significant. The negative squared term implies that ACT scores increase at a decreasing rate as SAT scores increase. This relationship between ACT scores and SAT scores is depicted in Figure 3.7.

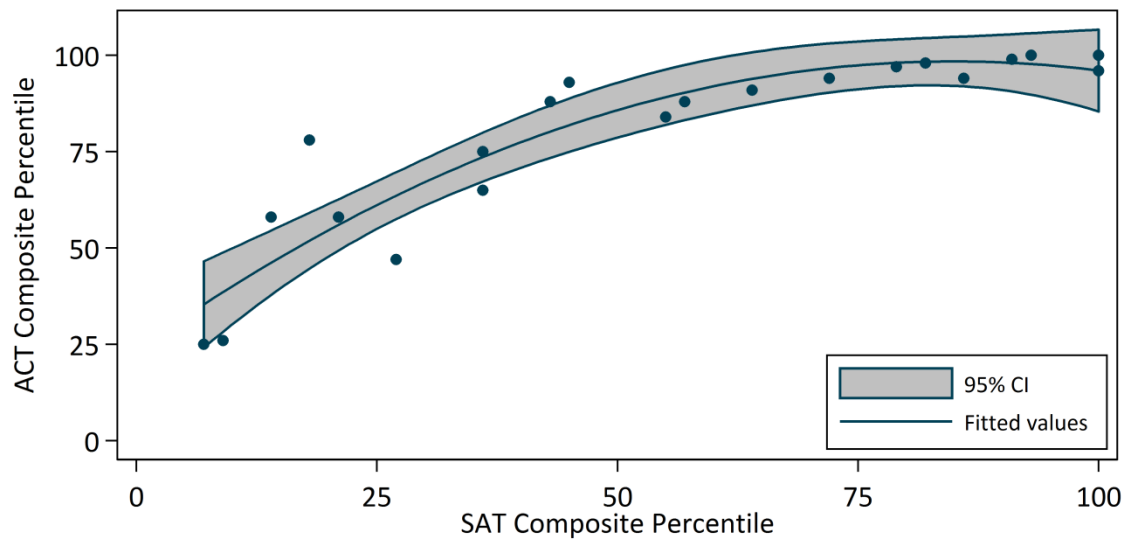
The R-squared of this model is 0.838, higher than the R-squared statistics of the other models discussed in this section. However, this is based on only 24 observations. As such, it is difficult to infer the true nature of the relationship between ACT and SAT performance at WISD if more students took both assessments.

Figure 3.6: SAT Predicting ACT, Regression Analysis

	COEFFICIENT	STANDARD ERROR
SAT Composite Percentile	1.718***	(0.284)
SAT Composite Percentile Squared	-0.010***	(0.003)
Constant	24.875***	(6.596)
Observations	24	
R-squared	0.838	

Notes: Coefficients are estimated using Ordinary Least Squares (OLS), using equation [1]. Numbers in parentheses denote standard errors. Asterisks denote statistical significance, as follows: *** p<0.01, ** p<0.05, * p<0.1.

Figure 3.7: SAT Predicting ACT Scatterplot



APPENDIX: CORRESPONDENCE BETWEEN ACTUAL AND PREDICTED SCORES

The following figure presents the predicted PLAN test scores for different values of PSAT test scores; the predicted ACT test scores for different values of PLAN test scores; the predicted ACT test scores for different values of PSAT test scores; and the predicted SAT test scores for different values of ACT test scores. These are computed using the results displayed in Figures 3.1, 3.3., 3.5, and 3.7, respectively.

For instance, we estimate that a student who scored in 50th percentile on the PSAT is expected to score in 67th percentile on the PLAN assessment. However, a student scoring in the 70th percentile of the PSAT assessment is expected to score in the 80th percentile of the PLAN assessment. This is due to the non-linear relationship between the two assessments.

It is also notable that, on average, regardless of ACT test score, all corresponding SAT predicted scores are an order of magnitude higher. For instance, the median ACT test score corresponds to the 86th percentile of SAT test scores within WISD, and all test scores above the 60th percentile on the ACT correspond to a test score that is greater than the 92nd percentile on the SAT.

Figure A.1: Percentile Score Correspondence between Actual and Predicted Scores (PSAT-PLAN; PLAN-ACT; PSAT-ACT; and ACT-SAT)

PSAT Actual Score	PLAN Predicted Score	PLAN Actual Score	ACT Predicted Score	PSAT Actual Score	ACT Predicted Score	ACT Actual Score	SAT Predicted Score
0	20	0	18	0	14	0	25
10	31	10	21	10	26	10	41
20	41	20	25	20	38	20	55
30	50	30	31	30	48	30	67
40	59	40	37	40	57	40	78
50	67	50	44	50	65	50	86
60	74	60	53	60	73	60	92
70	80	70	62	70	79	70	96
80	86	80	72	80	84	80	98
90	90	90	83	90	89	90	98
100	94	100	96	100	92	100	97

PROJECT EVALUATION FORM

Hanover Research is committed to providing a work product that meets or exceeds partner expectations. In keeping with that goal, we would like to hear your opinions regarding our reports. Feedback is critically important and serves as the strongest mechanism by which we tailor our research to your organization. When you have had a chance to evaluate this report, please take a moment to fill out the following questionnaire.

<http://www.hanoverresearch.com/evaluation/index.php>

CAVEAT

The publisher and authors have used their best efforts in preparing this brief. The publisher and authors make no representations or warranties with respect to the accuracy or completeness of the contents of this brief and specifically disclaim any implied warranties of fitness for a particular purpose. There are no warranties that extend beyond the descriptions contained in this paragraph. No warranty may be created or extended by representatives of Hanover Research or its marketing materials. The accuracy and completeness of the information provided herein and the opinions stated herein are not guaranteed or warranted to produce any particular results, and the advice and strategies contained herein may not be suitable for every partner. Neither the publisher nor the authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Moreover, Hanover Research is not engaged in rendering legal, accounting, or other professional services. Partners requiring such services are advised to consult an appropriate professional.



4401 Wilson Boulevard, Suite 400

Arlington, VA 22203

P 202.559.0500 F 866.808.6585

www.hanoverresearch.com