Duquesne University
Duquesne Scholarship Collection

Electronic Theses and Dissertations

Summer 8-10-2019

Score Improvement Comparison from PSAT to SAT Between two Schools

Salvatore Brusco

Follow this and additional works at: https://dsc.duq.edu/etd

Recommended Citation

Brusco, S. (2019). Score Improvement Comparison from PSAT to SAT Between two Schools (Master's thesis, Duquesne University). Retrieved from https://dsc.duq.edu/etd/1800

This Immediate Access is brought to you for free and open access by Duquesne Scholarship Collection. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Duquesne Scholarship Collection.

SCORE IMPROVEMENT COMPARISON FROM PSAT TO SAT BETWEEN TWO SCHOOLS

A Thesis

Submitted to the McAnulty College and Graduate School of Liberal Arts

Duquesne University

In partial fulfillment of the requirements for

the degree of Master of Science

By

Salvatore R. Brusco

May 2019

Copyright by

Salvatore R. Brusco

2019

SCORE IMPROVEMENT COMPARISON FROM PSAT TO SAT

BETWEEN TWO SCHOOLS

By

Salvatore R. Brusco

Approved May 22, 2019

John Kern, Ph.D., Associate Dean
Department of Mathematics and Computer Science

Frank D'Amico, Ph.D., Professor
Department of Mathematics and Computer Science

Stacey Levine, Ph.D., Department Chair
Department of Mathematics and Computer Science

James Swindal, Ph.D., Dean
McAnulty College and Graduate School of Liberal Arts

ABSTRACT

SCORE IMPROVEMENT COMPARISON FROM PSAT TO SAT

BETWEEN TWO SCHOOLS

By

Salvatore R. Brusco

May 2019

Thesis supervised by John C. Kern II, Ph.D.

The score a student earns on the PSAT is a predictor for what they earn on the SAT. Based on a student's PSAT score, we construct a Bayesian model to predict, with 95% certainty, what they will earn on their SAT. Furthermore, we explore differences in this prediction between two high schools in the same district; and ask whether the probability of improvement from the PSAT to the SAT is consistent between these two schools.

DEDICATION

I would like to dedicate this paper to everyone who encouraged me to complete this degree during the times it was hardest for me to persevere.

ACKNOWLEDGEMENT

I would like to thank John C. Kern II for his advisement through this process as well as supplying different ideas on how to approach this project. I would also like to thank the two anonymous schools for their data.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

1.1 Motivation

Most higher education institutions use the Scholastic Aptitude Test (SAT) or American College Testing (ACT) test as an indicator of how well a student did in high school. The Preliminary Scholastic Aptitude Test (PSAT) is supposed to be a predictor of performance on the SAT. The ACT Aspire or PLAN tests are analogous with the PSAT; however, the SAT has been around since 1926 and the ACT has been around since 1959, so this research will focus only on the ability of PSAT scores to predict SAT scores. Data collected for this project shows a strong correlation between a student's first SAT score and their previous PSAT score (see section 1.3). The goals of this research are to determine whether there is a difference between change in score of the PSAT to the SAT given two high schools in the same district, quantify an expected change in score from PSAT to SAT for the two schools, and find the probability for an increase in score from PSAT to SAT for both schools.

1.2 The Data

Since 2015, the SAT has two required parts: Evidence-Based Reading and Writing (ERW) and Mathematics. Each of these has a possible score range from 200 to 800 for a total possible score of 400 to 1600. PSAT scores and SAT scores were collected from students in two high schools in the same district. The students' first SAT score and their immediately prior PSAT score were gathered. Immediately prior is specified (instead of last) because on rare instances, students took the PSAT again after they took the SAT for the first time.

The two schools will remain anonymous throughout this analysis and will be referred to as School 1 and School 2. All data was gathered directly from the College Board website. Data from 183 students was gathered from both schools: 82 from School 1 and 101 from School 2. The graduation dates of students from School 1 range from 2017 to 2020 and the graduation dates of students from School 2 range from 2018 to 2020. The statistical software R will be used to analyze the data, beginning with an exploratory data analysis.

1.3 Exploratory Data Analysis

Figure A.1 shows scatter plots of the data, by school, for the total SAT versus the total PSAT scores. It appears that Schools 1 and 2 have more data points above the 45° line than below indicating students generally have an improved score on the SAT after taking the PSAT. There is a positive, linear association between the SAT and PSAT scores. Pearson correlation yields a value of 0.867 for School 1 and 0.865 for School 2.

Figure A.2 shows scatter plots of the data, by school, for the math section score of the SAT versus the math section score of the PSAT. There is a positive, linear association between the math section scores of the PSAT and the SAT. Pearson correlation yields values of 0.794 for School 1 and 0.819 for School 2.

Figure A.3 shows scatter plots of the data, by school, for the ERW section scores of the SAT versus the ERW section scores of the PSAT. There is a positive, linear association between the ERW section scores of the PSAT and the SAT. Pearson correlation yields values of 0.854 for School 1 and 0.828 for School 2.

Whether considering total score, math score, or ERW score, there are no large differences in SAT and PSAT correlation between the two schools that would lead to the conclusion that one set of scores is more correlated than another set of scores.

Table 1.1 shows the mean scores for School 1 and School 2 for the total score, math score, and ERW score from the PSAT and SAT. Table 1.1 also shows the changes between the PSAT and SAT for each of these categories. For the mean total score, students from School 1 earned 28 more points on average from the PSAT to the SAT and students from School 2 earned 38.8 more points on average. Similarly, the mean math score of students from School 1 increased 6.9 points while the mean math score of students from School 2 increased 21.6 points. Also, mean ERW scores of students

	School 1	School 2
Total Score	PSAT: 1004.8 SAT: 1032.8	PSAT: 1064.6 SAT: 1103.4
Change in Total Score	28.0	38.8
Math Score	PSAT: 493.8 SAT: 500.7	PSAT: 512.3 SAT: 533.9
Change in Math Score	6.9	21.6
ERW Score	PSAT: 511.0 SAT: 532.1	PSAT: 552.3 SAT: 569.5
Chang in ERW Score	21.1	17.2

Table 1.1: Mean scores for the PSAT and SAT for the total, math section, and ERW section scores for the two schools. Also displayed are the differences between SAT and PSAT mean scores.

from School 1 increased 21.1 points and the mean ERW scores of students from School 2 increased 17.2 points.

Figure A.4 shows boxplots of the individual scores by school (School 1 or School 2), by test area (total score, math section, or ERW section), and by test (PSAT or SAT). It is interesting to note the outliers from School 2 in the total PSAT and SAT scores and math PSAT and SAT scores; also the one outlier from School 1 in the math PSAT scores. Lastly, there were no outliers for the ERW section of either test for either school. Outliers will not be removed in this analysis.

Chapter 2: Bayesian Analysis Using Gibbs Sampling

2.1 Set-Up

Let the subscript i denote the ith student, Y_i be their score on the SAT, X_{1i} be their score on the PSAT, and X_{2i} be their school indicator. School 1 will be represented by $x_{2i} = 0$, and School 2 by $x_{2i} = 1$. Consider the following data model:

$$Y_i \sim N(b_0 + b_1 X_{1i} + b_2 X_{2i} + b_3 X_{1i} X_{2i}, \sigma^2) \tag{2.1}$$

The product of the n univariate normal densities for each Y_i gives the following likelihood function for b_0, b_1, b_2, b_3, and σ^2:

$$\begin{aligned} L(b_0, b_1, b_2, b_3, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_1 - \mu_{x_1})^2}{2\sigma^2}\right) \cdot \cdots \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_n - \mu_{x_n})^2}{2\sigma^2}\right) \\ &= \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left(-\frac{\sum (y_i - \mu_{x_i})^2}{2\sigma^2}\right). \end{aligned} \tag{2.2}$$

In this case, μ_{x_i} is the population mean that can be expressed as:

$$\mu_{x_i} = b_0 + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{1i} x_{2i}. \tag{2.3}$$

This data model will allow separate regression lines for the two schools to be expressed using a single regression equation. The joint prior density $\pi\left(b_0, b_1, b_2, b_3, \sigma^2\right)$ for the b_j and σ^2 variables can be expressed as

$$\pi\left(b_0, b_1, b_2, b_3, \sigma^2\right) = \pi\left(\sigma^2\right) \cdot \pi\left(b_0, b_1, b_2, b_3\right). \tag{2.4}$$

Let $\pi\left(\sigma^2\right) \propto \frac{1}{\sigma^2}$ (a reference prior [1]) and $\pi\left(b_0, b_1, b_2, b_3\right) \propto 1$ (a non-informative uniform prior). Then the joint posterior is found by multiplying the likelihood by the joint prior:

$$\begin{aligned} \pi\left(b_0, b_1, b_2, b_3, \sigma^2 | \vec{y}\right) &\propto \pi\left(\sigma^2\right) \cdot \pi\left(b_0, b_1, b_2, b_3\right) \cdot L\left(b_0, b_1, b_2, b_3, \sigma^2\right) \\ &\propto \frac{1}{\sigma^2} \cdot \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left(-\frac{\sum (y_i - \mu_{x_i})^2}{2\sigma^2}\right) \\ &\propto \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}+1} \exp\left(-\frac{\sum (y_i - \mu_{x_i})^2}{2\sigma^2}\right). \end{aligned} \tag{2.5}$$

Expanding the $(y_i - \mu_{x_i})$ term in 2.5 gives:

$$y_i^2 - 2y_i(b_0 + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i}) + (b_0 + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i})^2. \tag{2.6}$$

For simplicity, let $x_{1i}x_{2i} = x_{3i}$. Expanding 2.6 then gives:

$$\begin{aligned} y_i^2 - 2b_0y_i - 2b_1x_{1i}y_i - 2b_2x_{2i}y_i - 2b_3x_{3i}y_i + b_0^2 + b_1^2x_{1i}^2 + b_2^2x_{2i}^2 + b_3^2x_{3i}^2 \\ + 2b_0b_1x_{1i} + 2b_0b_2x_{2i} + 2b_0b_3x_{3i} + 2b_1b_2x_{1i}x_{2i} + 2b_1b_3x_{1i}x_{3i} + 2b_2b_3x_{2i}x_{3i}. \end{aligned} \tag{2.7}$$

We apply the summation to 2.5 and complete the square in b_0 to yield:

$$\begin{aligned} &\sum b_0^2 - 2b_0\left(\sum y_i - b_1\sum x_{1i} - b_2\sum x_{2i} - b_3\sum x_{3i}\right) \\ \Rightarrow &n\left(b_0^2 - 2b_0\frac{\sum y_i - b_1\sum x_{1i} - b_2\sum x_{2i} - b_3\sum x_{3i}}{n}\right) \\ \Rightarrow &n\left(b_0 - \frac{\sum y_i - b_1\sum x_{1i} - b_2\sum x_{2i} - b_3\sum x_{3i}}{n}\right)^2. \end{aligned} \tag{2.8}$$

The form of this completed square and its place in 2.5 implies the full conditional density of b_0 is normal:

$$\begin{aligned} &b_0|b_1, b_2, b_3, \sigma^2, \vec{y}, \vec{x_1}, \vec{x_2} \\ &\sim N\left(\frac{\sum y_i - b_1\sum x_{1i} - b_2\sum x_{2i} - b_3\sum x_{3i}}{n}, \frac{\sigma^2}{n}\right). \end{aligned} \tag{2.9}$$

In a similar way, we apply the summation to 2.7 and complete the square in b_1:

$$\begin{aligned} &b_1^2\sum x_{1i}^2 - 2b_1\left(\sum x_{1i}y_i - b_0\sum x_{1i} - b_2\sum x_{1i}x_{2i} - b_3\sum x_{1i}x_{3i}\right) \\ \Rightarrow &\sum x_{1i}^2\left(b_1^2 - 2b_1\frac{\sum x_{1i}y_i - b_0\sum x_{1i} - b_2\sum x_{1i}x_{2i} - b_3\sum x_{1i}x_{3i}}{\sum x_{1i}^2}\right) \\ \Rightarrow &\sum x_{1i}^2\left(b_1 - \frac{\sum x_{1i}y_i - b_0\sum x_{1i} - b_2\sum x_{1i}x_{2i} - b_3\sum x_{1i}x_{3i}}{\sum x_{1i}^2}\right)^2. \end{aligned} \tag{2.10}$$

This gives the full conditional for b_1 as normal:

$$\begin{aligned} &b_1|b_0, b_2, b_3, \sigma^2, \vec{y}, \vec{x_1}, \vec{x_2} \\ &\sim N\left(\frac{\sum x_{1i}y_i - b_0\sum x_{1i} - b_2\sum x_{1i}x_{2i} - b_3\sum x_{1i}x_{3i}}{\sum x_{1i}^2}, \frac{\sigma^2}{\sum x_{1i}^2}\right). \end{aligned} \tag{2.11}$$

Similar computations yield normal full-conditional densities for b_2 and b_3:

$$\begin{aligned} &b_2|b_0, b_1, b_3, \sigma^2, \vec{y}, \vec{x_1}, \vec{x_2} \\ &\sim N\left(\frac{\sum x_{2i}y_i - b_0\sum x_{2i} - b_1\sum x_{1i}x_{2i} - b_3\sum x_{2i}x_{3i}}{\sum x_{2i}^2}, \frac{\sigma^2}{\sum x_{2i}^2}\right) \end{aligned} \tag{2.12}$$

$$b_3|b_0, b_1, b_2, \sigma^2, \vec{y}, \vec{x_1}, \vec{x_2}$$
$$\sim N\left(\frac{\sum x_{3i}y_i - b_0\sum x_{3i} - b_1\sum x_{1i}x_{3i} - b_2\sum x_{2i}x_{3i}}{\sum x_{3i}^2}, \frac{\sigma^2}{\sum x_{3i}^2}\right). \tag{2.13}$$

For σ^2, referencing equation 2.5, the full conditional can be recognized as an inverse gamma density:

$$\sigma^2|b_0, b_1, b_2, b_3, \vec{y}, \vec{x_1}, \vec{x_2}$$
$$\sim IG\Bigg(\frac{n}{2}, \sum\left(y_i^2 - 2y_i(b_0 + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i})\right. \tag{2.14}$$
$$\left.+ (b_0 + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i})^2\right)\Bigg).$$

2.2 Gibbs Sampling

Having established full conditional densities for b_0, b_1, b_2, b_3, and σ^2, Gibbs sampling [2] is used to generate 10,000 posterior realizations of each b_j and of σ^2. All b_j were initialized at 1, and σ^2 was initialized at the variance of all PSAT scores. A lag of 800 was used to ensure that there was no relation between consecutive realizations and the first 799 eliminated from burn-in to ensure that the initialization values did not affect the inference. This burn-in of 799 is based on the lag of 800, which is why `burn = 0` in the Gibbs sampling code below.

```
N <- 10000
b0 <- 1
b1 <- 1
b2 <- 1
b3 <- 1
sigSquared <- var(x1)
n <- length(x1)
burn <- 0
lag <- 800
b0.s <- NULL
b1.s <- NULL
b2.s <- NULL
b3.s <- NULL
sigSquared.s <- NULL
for (i in 1:(burn + N * lag)) {
  sigSquared <- 1/rgamma(1,n/2,sum(y^2-2*y*(b0+b1*x1+b2*x2+b3*x3)
      +(b0+b1*x1+b2*x2+b3*x3)^2)/2)
  b0 <- rnorm(1,(ySum-b1*x1Sum-b2*x2Sum-b3*x3Sum)/n,sqrt(sigSquared/n))
  b1 <- rnorm(1,(x1ySum-b0*x1Sum-b2*x1x2Sum-b3*x1x3Sum)/x1SquaredSum,
      sqrt(sigSquared/x1SquaredSum))
```

```
  b2 <- rnorm(1,(x2ySum-b0*x2Sum-b1*x1x2Sum-b3*x2x3Sum)/x2SquaredSum,
      sqrt(sigSquared/x2SquaredSum))
  b3 <- rnorm(1,(x3ySum-b0*x3Sum-b1*x1x3Sum-b2*x2x3Sum)/x3SquaredSum,
          sqrt(sigSquared/x3SquaredSum))
  if (i > burn - 1 & i%%lag == 0) {
    sigSquared.s <- c(sigSquared.s,sigSquared)
    b0.s <- c(b0.s,b0)
    b1.s <- c(b1.s,b1)
    b2.s <- c(b2.s,b2)
    b3.s <- c(b3.s,b3)
  }
}
```

Figure A.5 shows the trace plots of the b_j and σ^2 variables. Without a burn-in, it is possible to incorporate values sampled early in the chain before the chain has converged to its stationary distribution which in turn would bias posterior distribution estimates. Since figure A.5 shows there is convergence to a stationary distribution, the burn-in of 799 is sufficient. The lag of 800 was selected based on autocorreltion plots that showed negligible autocorrelation for all variables at lag 800. Figure A.6 shows the autocorrelation plots of the variable realizations having saved only every 800th value; notice the immediate drop to negligible autocorrelation at lag 1 for all variables. Similar results can be seen for the math score data in figures A.7 and A.8 and for the ERW score data in figures A.9 and A.10.

2.3 Posterior Mean Regression Line

Subsequently, a vector of 300 equally spaced x_1-values (PSAT scores) were generated between the minimum and maximum PSAT score. Since $x_2 = 0$ for School 1 and $x_2 = 1$ for School 2, the following two formulas yield generated posterior regression lines.

$$\hat{y} = b_0 + b_1 x_1 \tag{2.15}$$

$$\hat{y} = (b_0 + b_2) + (b_1 + b_3) x_1. \tag{2.16}$$

Equation 2.15 is for School 1 ($x_2 = 0$) and equation 2.16 is for School 2 ($x_2 = 1$).

Using the 10,000 b_j values obtained from the above Gibbs sampling code, 300 $\hat{y}$ (SAT scores) realizations (once for each generated x_1 value) were generated $10,000$ times, once for each set of b_j. This resulted in two $10,000$ by 300 matrices for predicted values of $\hat{y}$ (one matrix for each school). The following code finds the y-coordinates of all of the posterior regression lines for each of the discretized x_1-coordinates.

```
xx <- seq(min(x1),max(x1),length = 300)

PMRL1 <- NULL
for (i in 1:10000) {
  y1 <- b0.s[i] + b1.s[i]*xx
  PMRL1 <- rbind(PMRL1,y1)
}

PMRL2 <- NULL
for (i in 1:10000) {
  y2 <- b0.s[i] + b2.s[i] + (b1.s[i] + b3.s[i])*xx
  PMRL2 <- rbind(PMRL2,y2)
}
```

The `apply` function in R was used to find the mean, 2.5% quantile, and 97.5% quantile of the y-coordinates of the posterior regression lines. The following code generates the posterior mean regression line and corresponding 95% credible interval for both schools.

```
points(xx,apply(PMRL1,2,quantile,probs=.025,na.rm=TRUE),pch=".")
points(xx,apply(PMRL1,2,quantile,probs=.975,na.rm=TRUE),pch=".")
points(xx,apply(PMRL1,2,"mean"),pch=20)

points(xx,apply(PMRL2,2,quantile,probs=.025,na.rm=TRUE),pch="."
    ,col="red")
points(xx,apply(PMRL2,2,quantile,probs=.975,na.rm=TRUE),pch="."
    ,col="red")
points(xx,apply(PMRL2,2,"mean"),pch=20,col="red")
```

2.4 Posterior Predictive Distribution

The same vector of 300 equally spaced x_1-values (PSAT scores) was used for the posterior predictive distribution. Using the 10,000 b_j and σ^2 values obtained from the above Gibbs sampling code, 300 normal realizations (once for each generated x_1 value) were generated $10,000$ times, once for each set of b_j and σ^2. This resulted in a $10,000$ by 300 matrix of predicted y-values. This process for generating predicted y-values is reflected in the following code.

```
yValSchool1 <- NULL
for (i in 1:10000) {
  draw <- rnorm(length(xx),b0.s[i] + b1.s[i]*xx,sqrt(sigSquared.s[i]))
  yValSchool1 <- rbind(yValSchool1,draw)
}

yValSchool2 <- NULL
for (i in 1:10000) {
  draw <- rnorm(length(xx),b0.s[i] + b2.s[i] + (b1.s[i] + b3.s[i])
    * xx,sqrt(sigSquared.s[i]))
  yValSchool2 <- rbind(yValSchool2,draw)
}
```

Using the output from the above code, the posterior predictive distribution of each column for each school was generated: the middle 95% credible interval and mean. The following code finds the means and quantiles.

```
school1PPD <- NULL
for (i  in 1:300) {
  a <- quantile(yValSchool1[,i],.025)
  b <- mean(yValSchool1[,i])
  c <- quantile(yValSchool1[,i],.975)
  school1PPD <- cbind(school1PPD,c(a,b,c))
}

school2PPD <- NULL
for (i in 1:300) {
  a <- quantile(yValSchool2[,i],.025)
  b <- mean(yValSchool2[,i])
  c <- quantile(yValSchool2[,i],.975)
  school2PPD <- cbind(school2PPD,c(a,b,c))
}
```

2.5 Results

Figure A.11 shows the histograms of the 10,000 draws from the marginal posterior distribution of each of b_0, b_1, b_2, b_3, and σ^2. For b_2 and b_3, 0 is very close to the center of the histogram, and represents a lack of evidence that the population regression lines for School 1 and School 2 differ.

Figure A.12 shows the data with superposed posterior mean regression lines for both schools. Note that neither of the posterior mean regression lines intersect in the domain of the plot and School 2's total scores are, on average, always greater than School 1's total scores. Shown in figure A.13 is a plot of 500 posterior regression lines for each school. These 500 are representative of all 10,000 posterior regression lines obtained from the Gibbs sampling procedure. Figure A.14 displays the graph of the 95% credible interval for the true posterior mean regression line. The mean posterior regression lines for School 1 and School 2 are the same as the posterior mean regression lines for School 1 and School 2 from figure A.12. Figure A.15 shows the generated posterior predictive distributions for both schools. The added blue, dashed line in figures A.12, A.14, and A.15 is the blue, dashed 45° line that represents equality of PSAT score and SAT score.

Chapter 3: Individual Sections of the PSAT and SAT Exams

3.1 Mathematics Section

The same code used in the previous chapter to analyze the total scores of the PSAT and the SAT is now used to analyze the mathematics scores of the PSAT and SAT to obtain posterior realizations of b_0, b_1, b_2, b_3, and σ^2 based only on the mathematics section scores. The values for the sets of b_j and σ^2 are still fairly normal, as seen from the marginal posterior histograms in figure A.16. In b_0, b_2, and b_3, 0 is contained within the 95% credible interval. This observation will be referenced again in section 4.2.

Figure A.17 shows nearly parallel posterior mean regression lines superposed over the scatter plot of SAT versus PSAT math scores for both schools. Note that the majority of these posterior mean regression lines are above the 45° line showing meaning that scores on the math section of the SAT are generally higher than those of the PSAT. Figure A.18 (as in figure A.13) shows 500 posterior regression lines for each school. These are representative of all 10,000 for both schools. Figure A.19 shows the mean posterior regression lines and the 95% credible intervals for the true posterior mean regression lines for both schools. It is interesting to note that the 45° line is completely contained in School 1's credible interval, but only partially contained in School 2's credible interval. Figure A.20 shows the posterior predictive distribution for the math sections of the PSAT and the SAT. School 1's mean posterior regression line is very close to the blue, dashed 45° line.

3.2 Evidence-Based Reading and Writing

The same code was applied to the ERW scores of the PSAT and SAT. Figure A.21 shows histograms of the posterior realizations of the b_j and σ^2 variables. Once again, 0 is contained within the 95% credible interval for b_0, b_2, and b_3, (see section 4.2).

Figure A.22 shows the posterior mean regression lines for both schools. These posterior mean regression lines intersect much lower (in reference to PSAT scores) than the posterior mean regression lines of the total score or the math score showing that School 1 starts out with higher scores, but School 2 ends with higher scores. Figure A.23 shows a representative 500 of the 10,000 posterior regression lines for the ERW scores for each school. Both of these graphs look very similar. Figure A.24 shows the mean posterior regression lines and the 95% credible intervals for the true posterior mean regression lines for both schools. These two graphs seem to have very similar credible interval sizes (see section 4.1). Figure A.25 shows the posterior predictive distribution for the ERW scores of the PSAT and the SAT for both schools. School 2 has some fairly distant outliers.

Chapter 4: Conclusion

4.1 Expected SAT Scores

Students often wonder, "What score can I expect to get on the SAT?" The credible interval widths from the posterior predictive distributions of the total score can be used to begin to answer this question. According to the mean distance between the posterior mean scores and the prediction intervals' upper and lower quantiles, a student from School 1 could expect their SAT score to be within 155 of their PSAT score with 95% credibility. Similarly for School 2, a student could expect to get a total SAT score within 154 of their total PSAT score with 95% credibility. This was calculated in R and can be estimated from the vertical distances between the 2.5% and 97.5% quantiles shown in figure A.26. These distances range from around 300 to 320. The average distance is about 310, which yields the ± 155 for students from School 1 (or ± 154 for School 2) margin of error.

The expected change in SAT score is now computed, be it positive or negative, as compared to the student's PSAT score. This is obtained from the posterior mean regression line. For School 1 (using the means of b_0 and b_1), this is $SAT = 0.889 * PSAT + 139.9$. The change is demonstrated by $SAT - PSAT = PSAT(0.889 - 1) + 139.9$. Combining this and the credible intervals, it can be said with 95% credibility that a student from School 1 that has taken the PSAT can expect to get a change in score of $(-0.111 * PSAT + 139.9) \pm 155$ on their SAT. In a similar manner, it can be said with 95% credibility that a student from School 2 that has taken the PSAT can expect to get a change in score of $(-0.128 * PSAT + 174.6) \pm 154$ on their SAT.

When examining (in the same way) the math section scores of School 1, it is expected that the SAT score will be within 110 of the PSAT score and students from School 2 will be within 109; this conclusion is based on the interval widths show in figure A.27. On the ERW section, for both schools, it is expected that the SAT score will be within 88 of the PSAT score (see figure A.28).

The credible intervals for the change in PSAT to SAT math and ERW scores can be calculated in the same way as the credible intervals for the change in total PSAT to total SAT scores. For the math section, given their corresponding score on the PSAT, a student from School 1 can say with 95% credibility that their score would change within the interval of $(-0.083 * PSAT + 47.9) \pm 110$. Similarly for School 2, a student that has taken the math section of the PSAT could expect get a change in score of $(-0.139 * PSAT + 92.7) \pm 109$ with 95% credibility. For the ERW section, a student from School 1 can be 95% certain that, given their last PSAT score, they would get a change in score of $(-0.258 * PSAT + 153.1) \pm 88$ on their SAT. Similarly for School 2, a student that has taken the ERW section of the PSAT can have a prediction of $(-0.195 * PSAT + 125.0) \pm 88$ change in score that is 95% credible.

4.2 Significance

The question posed now is, "Is there a significant difference between the two schools?" This question is answered through separate examination of credible intervals for b_2 and b_3 for the total score, math score, and ERW score. For b_2, the coefficient controlling the change in y-intercept between the two schools, a credible interval of $(-124.131, 193.942)$ is obtained for total scores. For b_3, the coefficient controlling for the change in slope, a credible interval of $(-0.166, 0.136)$ is obtained for total scores. Both of these intervals include 0. This leads to the conclusion that there is no significant difference between change in total score between the two schools. Table 4.1 summarizes the 95% credible intervals for the b_2 and b_3 parameters as estimated separately for the total scores, math only scores, and ERW only scores. Note that 0 is contained in all of these intervals.

	b_2	b_3
Total	$(-124.131, 193.942)$	$(-0.166, 0.136)$
Math	$(-53.310, 144.310)$	$(-0.251, 0.138)$
ERW	$(-107.303, 51.369)$	$(-0.083, 0.209)$

Table 4.1: The 95% credible intervals for b_2 and b_3 demonstrating inclusion of 0; and therefore, no significant difference between the schools.

4.3 Probability of Improvement

Figures A.29, A.30, and A.31 depict the probability of a student getting a better score on their SAT than on their PSAT. Figure A.29 shows the probabilities for the improvement of total score. Figure A.30 shows the probabilities for improvement on the math score. Figure A.31 shows the probabilities for improvement on the ERW score. Some things that are important to note, based upon figure A.29, is that students from School 2 with low total PSAT scores have about a 7% greater chance of improving their score than students from School 1. This difference shrinks to an approximate 3% greater chance at high PSAT scores. For math section scores, students from School 2 with lower PSAT scores have about a 13% better chance of getting a better score on the SAT math section than students from School 1. This difference approaches 3% as PSAT scores approach 700. It is interesting to look at the probability of improvement on the ERW section. For lower scores on the ERW section of the PSAT, students from School 1 have a slightly greater probability of improvement than students from School 2. However, this changes for scores around 700 on the ERW section of the PSAT, where students from School 2 have about a 14% greater probability of improvement. These probabilities are outlined in table 4.2.

	Low Score	Medium Score	High Score
Total Score	School 1: 82% School 2: 89%	School 1: 65% School 2: 75%	School 1: 44% School 2: 47%
Math Score	School 1: 67% School 2: 80%	School 1: 55% School 2: 68%	School 1: 45% School 2: 48%
ERW Score	School 1: 98% School 2: 95%	School 1: 70% School 2: 73%	School 1: 25% School 2: 39%

Table 4.2: Probability of improvement for each school given their score on the PSAT. A low score on the math or ERW sections is 300. A medium score on the math or ERW section is 500. A high score on the math or ERW section is 700. A low total score is 600. A medium total score is 1000. A high total score is 1400.

These probabilities of improvement lead to the conclusion that, in general, students from School 2 have a modestly higher chance of improvement.

4.4 Limitations

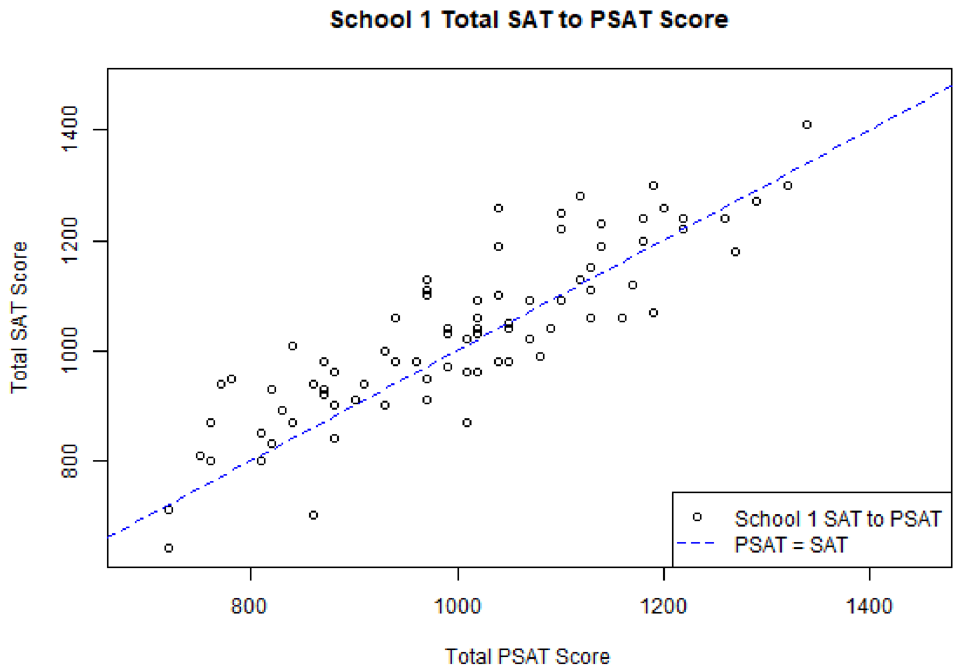
As far as demographics are concerned, School 2 is about 3 times as big as School 1. Also School 1 resides in an area of lower economic status. Both of these factors would normally be indicative of favoritism to School 2, as far as improvement in scores is concerned and could explain the slightly higher improvements and chances of improvement seen for School 2. Further studies may research the correlation of score improvement with demographics.

References

1. Berger, J. O.; Bernardo, J. M.; Sun, D. "The Formal Definition of Reference Priors". *The Annals of Statistics* **2009**, *37*, 905–38.

2. Cascella, G.; George., E. I. "Explaining the Gibbs Sampler". *The American Statistician* **1992**, *46*, 167–74.

Appendix A: Figures



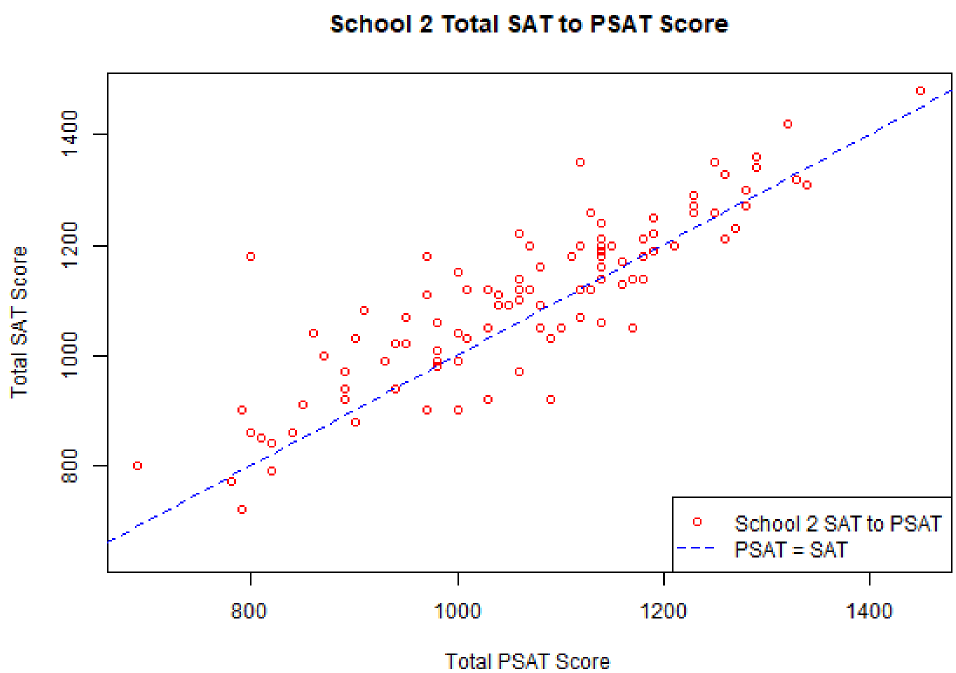


Figure A.1: Scatter plots of the total SAT score versus the total PSAT score separated by school. School 1 is on the top. School 2 is on the bottom. For both plots, the blue, dashed line is the 45° line representing equality between total PSAT score and total SAT score.

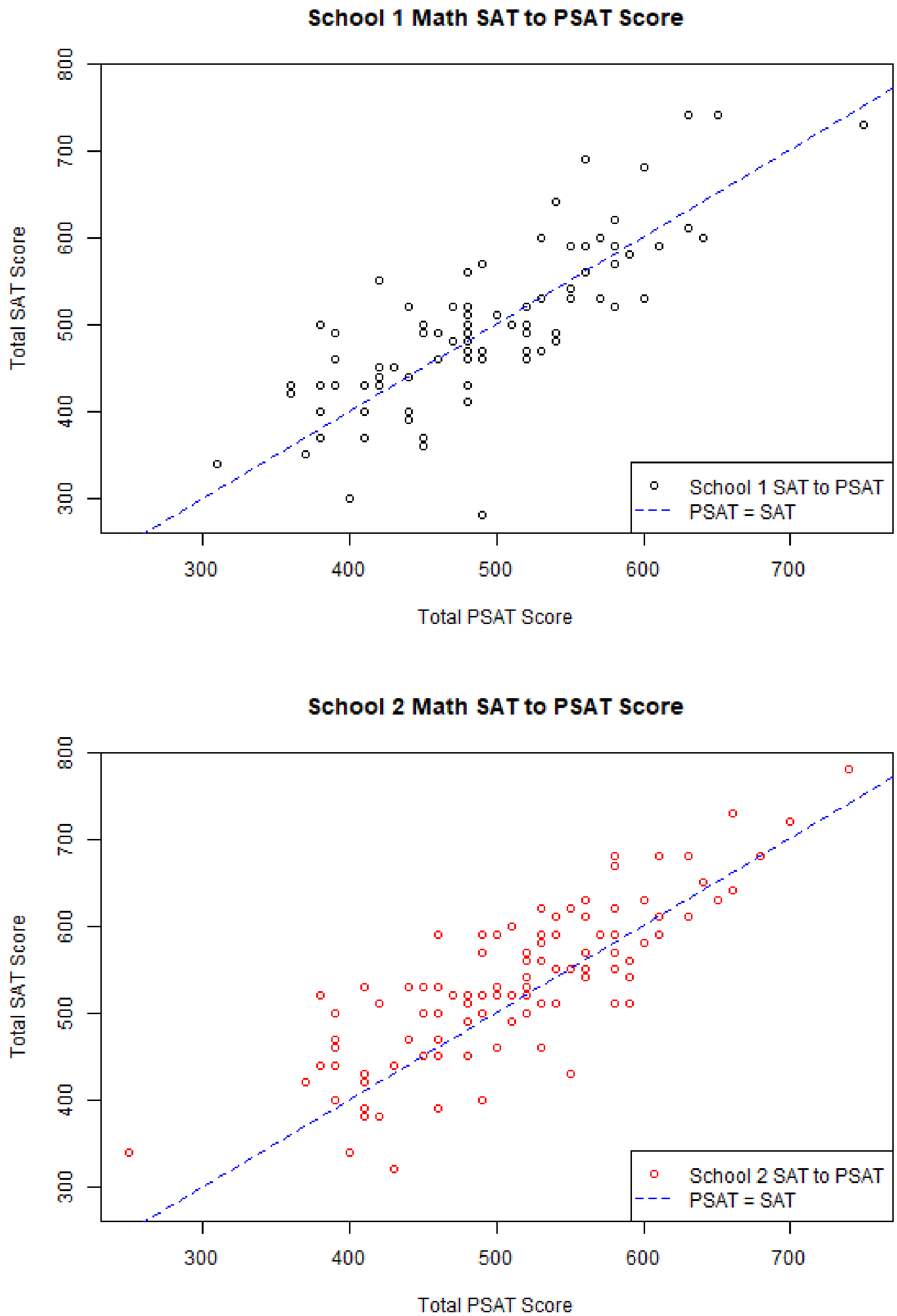
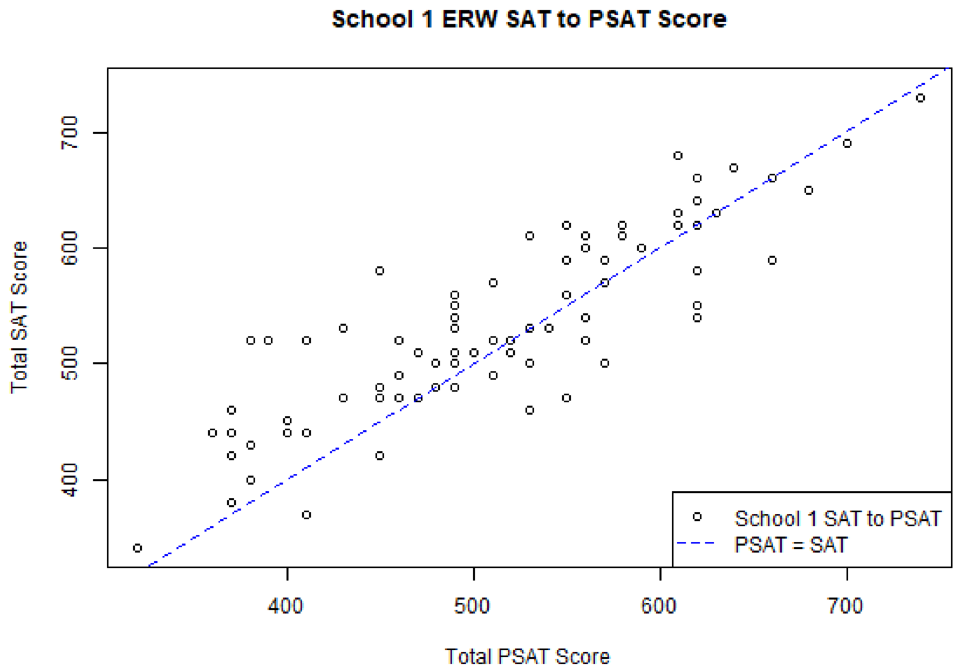


Figure A.2: Scatter plots of the math section SAT score versus the math section PSAT score separated by school. School 1 is on the top. School 2 is on the bottom. For both plots, the blue, dashed line is the 45° line representing equality between total PSAT score and total SAT score.



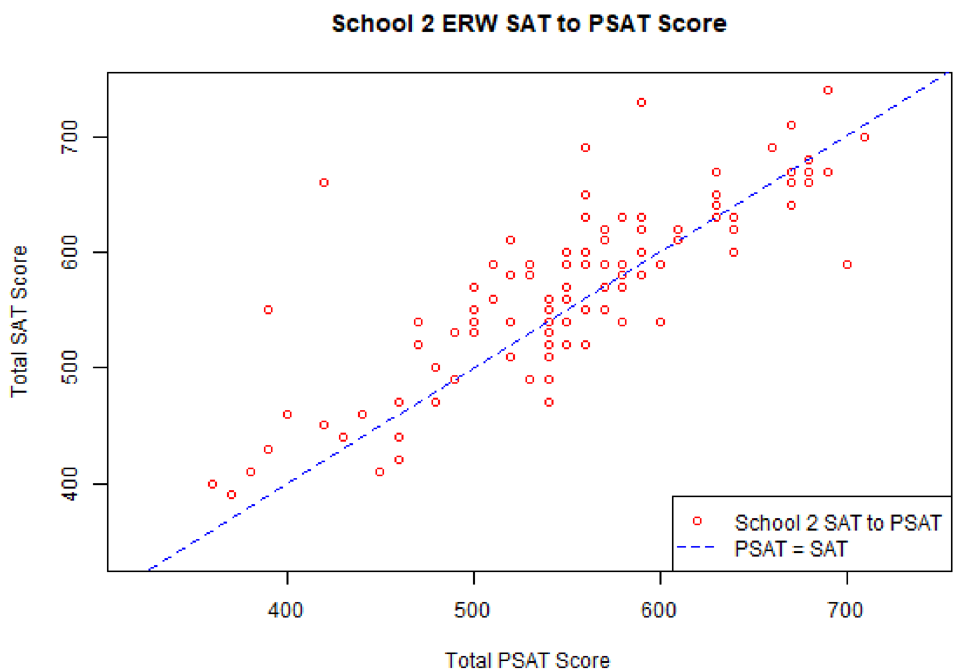


Figure A.3: Scatter plots of the ERW section SAT score versus the ERW section PSAT score separated by school. School 1 is on the top. School 2 is on the bottom. For both plots, the blue, dashed line is the 45° line representing equality between total PSAT score and total SAT score.

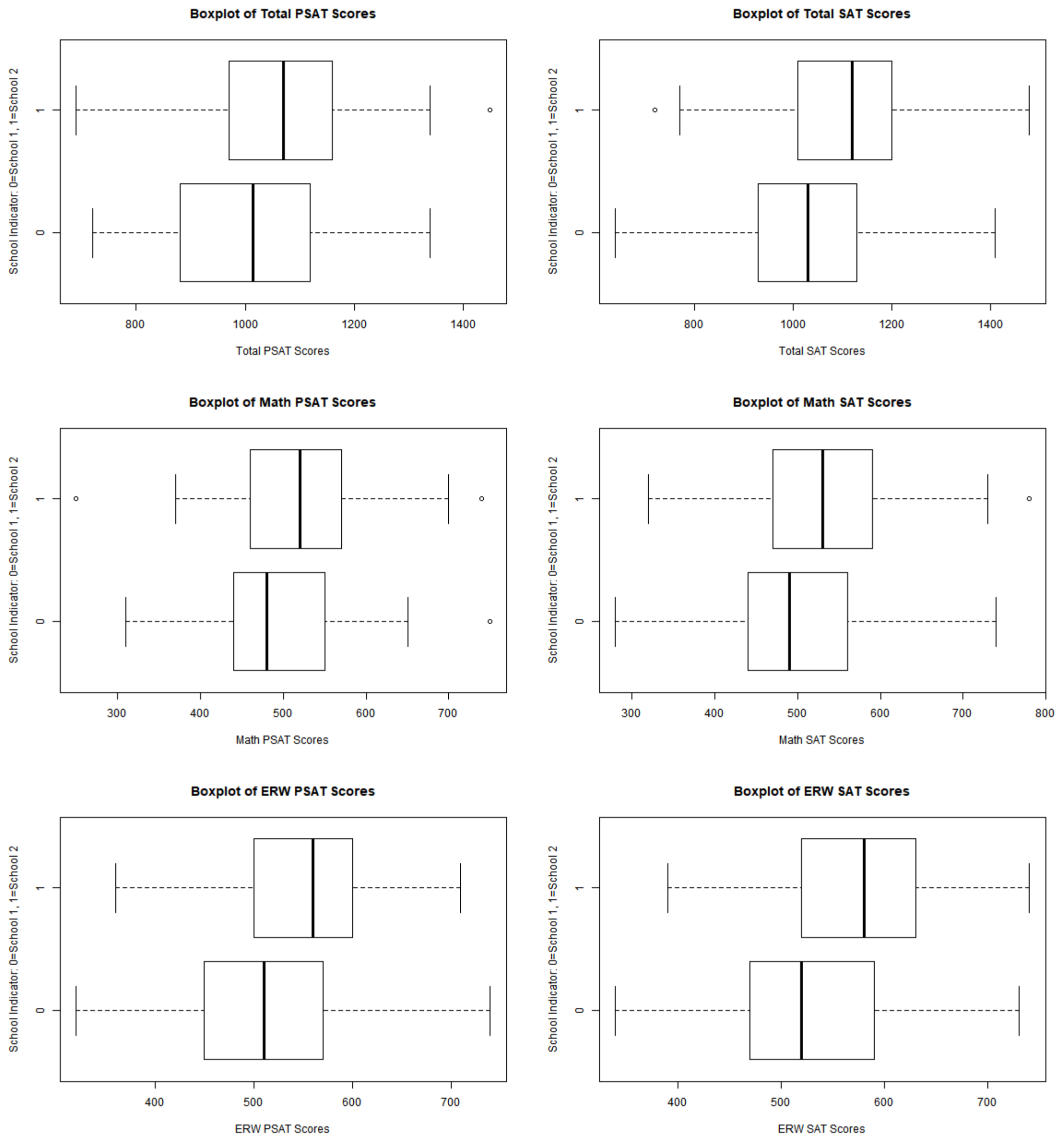


Figure A.4: Boxplots of the data by School (School 1 or School 2), test area (total score, math section, or ERW section), and test (PSAT or SAT). School 1 is the bottom boxplot and School 2 is the top boxplot.

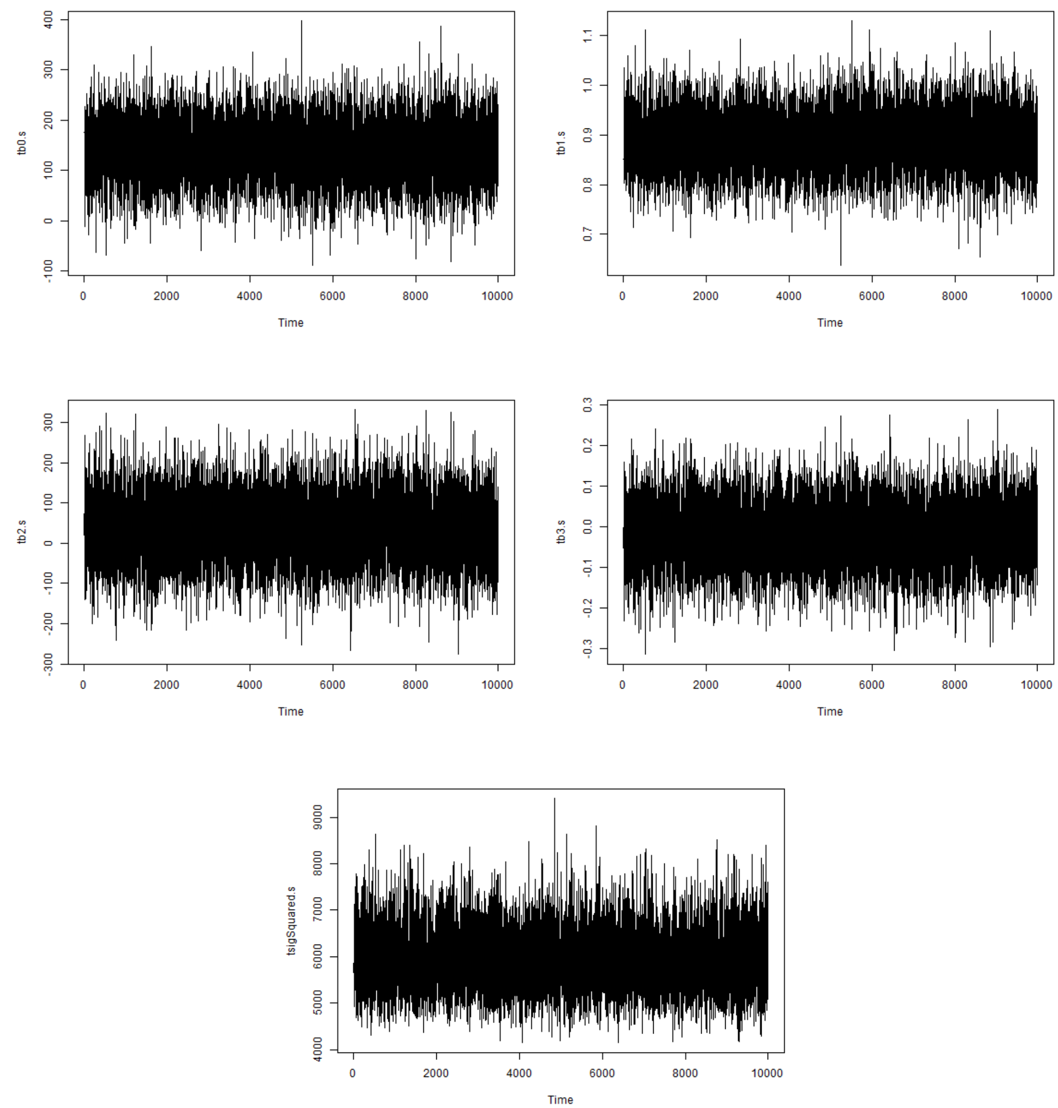


Figure A.5: Trace plots of b_0, b_1, b_2, b_3, and σ^2 for the posterior regression lines of the total PSAT and SAT scores.

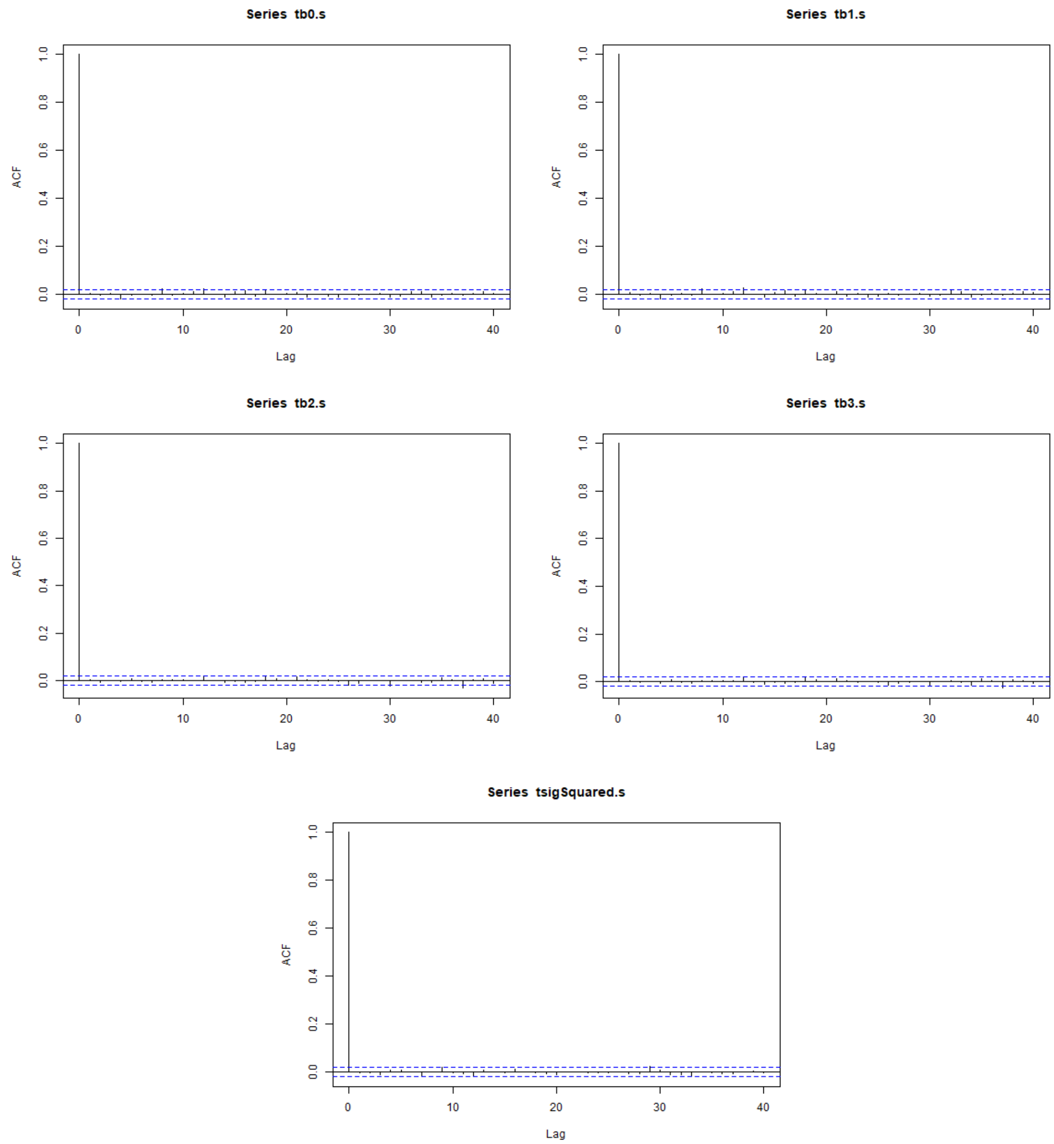


Figure A.6: Autocorrelation function of b_0, b_1, b_2, b_3, and σ^2 for the posterior regression lines of the total PSAT and SAT scores.

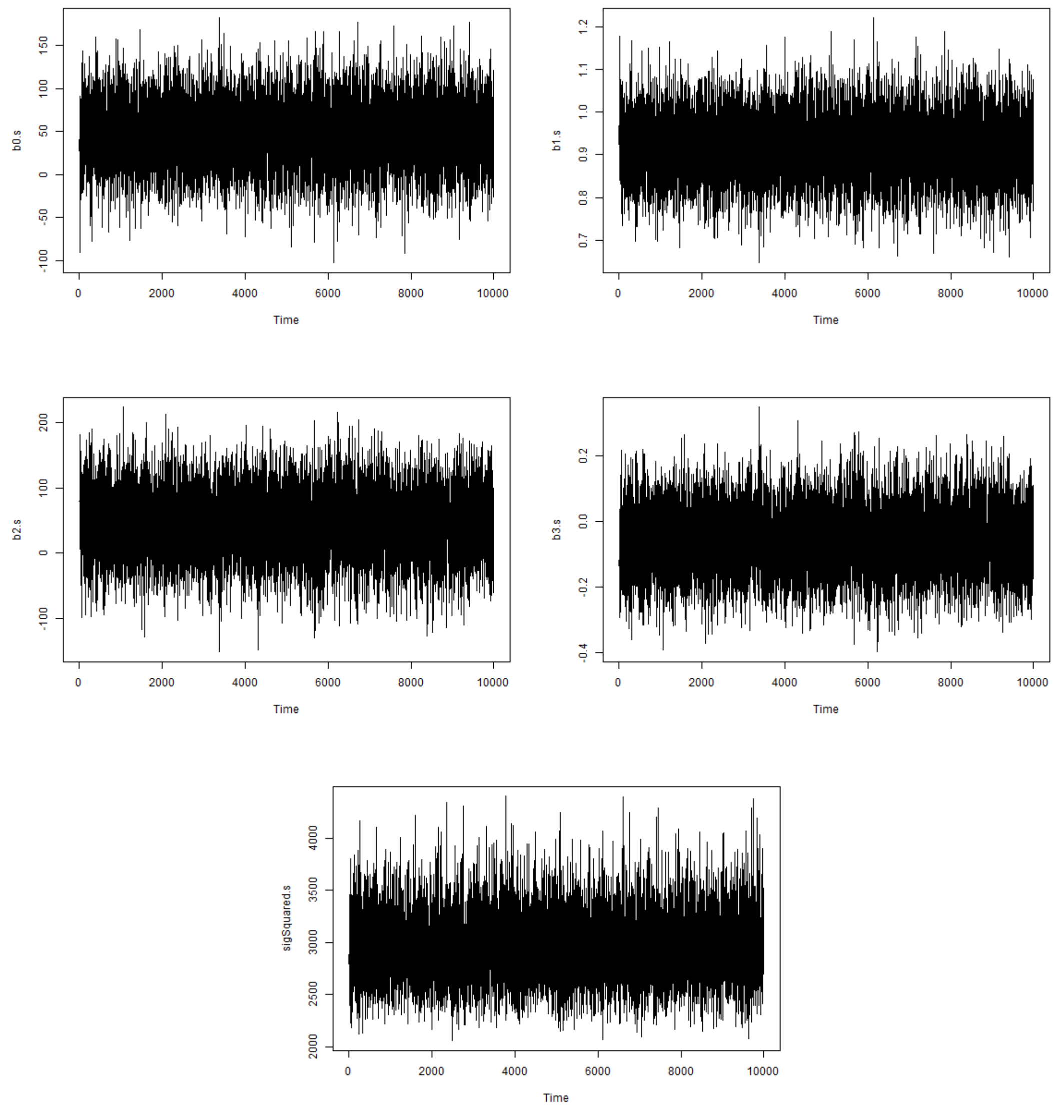


Figure A.7: Trace plots of b_0, b_1, b_2, b_3, and σ^2 for the posterior regression lines of the mathematics section scores of the PSAT and SAT.

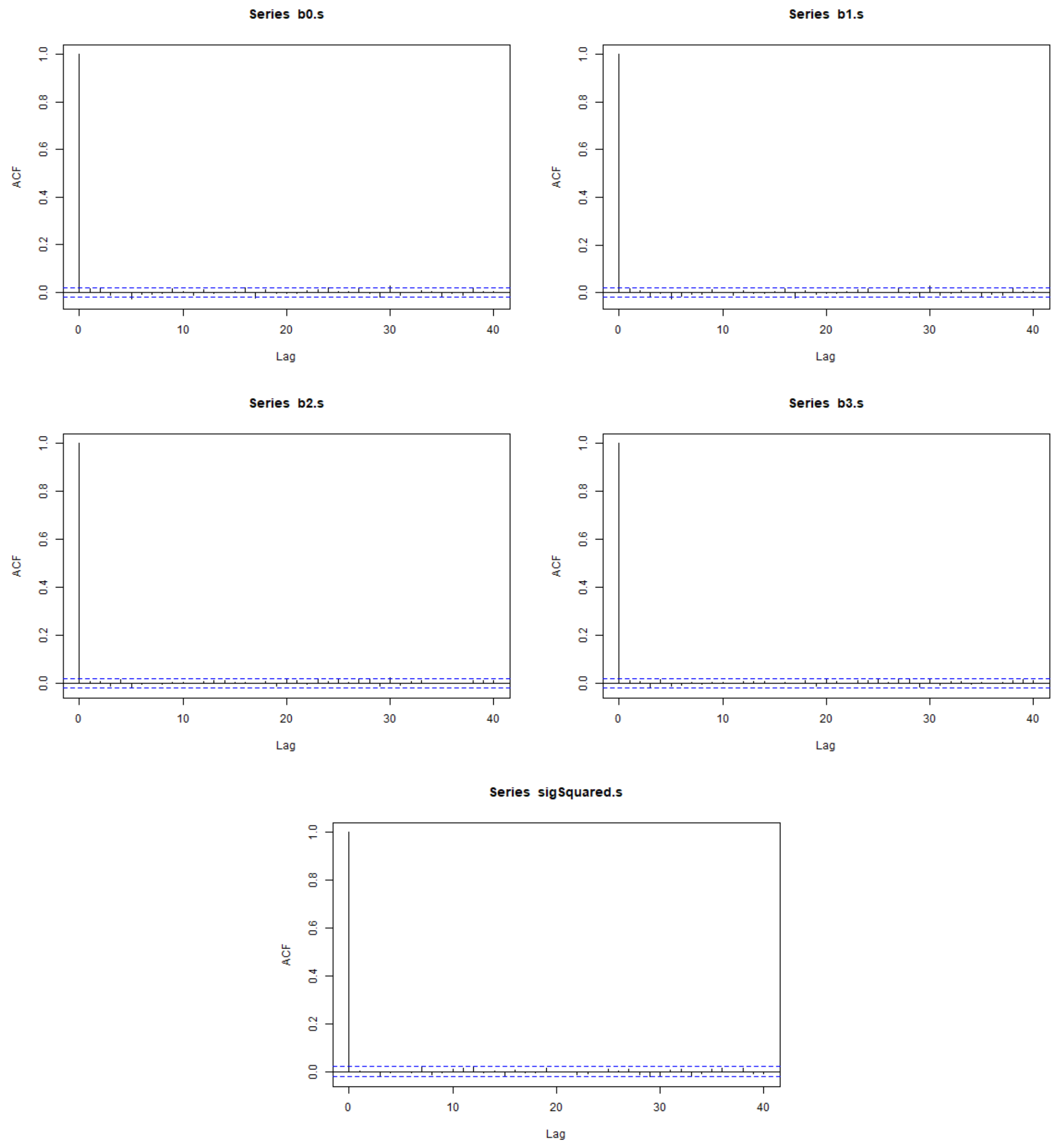


Figure A.8: Autocorrelation function of b_0, b_1, b_2, b_3, and σ^2 for the posterior regression lines of the mathematics section scores of the PSAT and SAT.

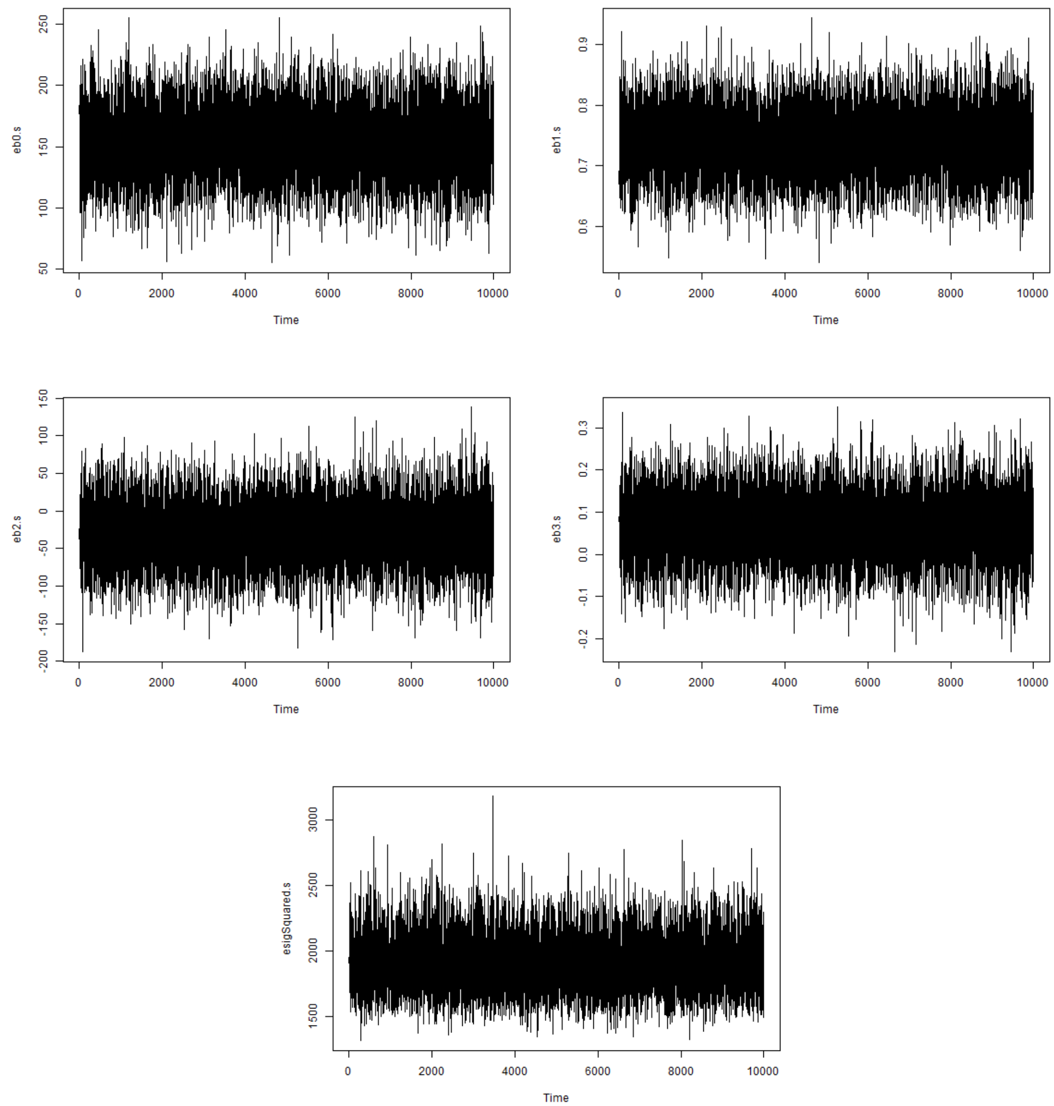


Figure A.9: Trace plots of b_0, b_1, b_2, b_3, and σ^2 for the posterior regression lines of the ERW scores of the PSAT and SAT.

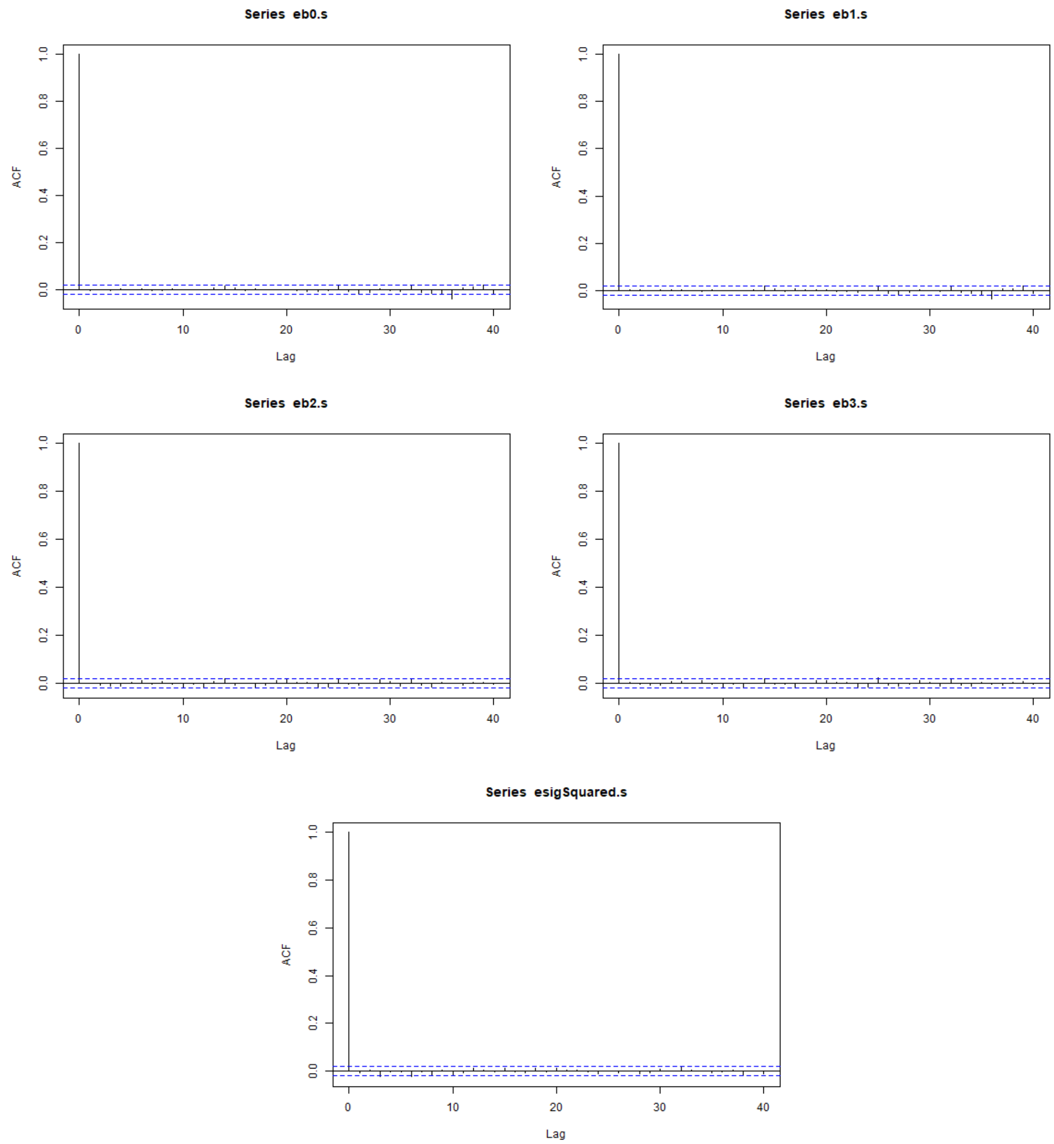


Figure A.10: Autocorrelation function of b_0, b_1, b_2, b_3, and σ^2 for the posterior regression lines of the ERW scores of the PSAT and SAT.

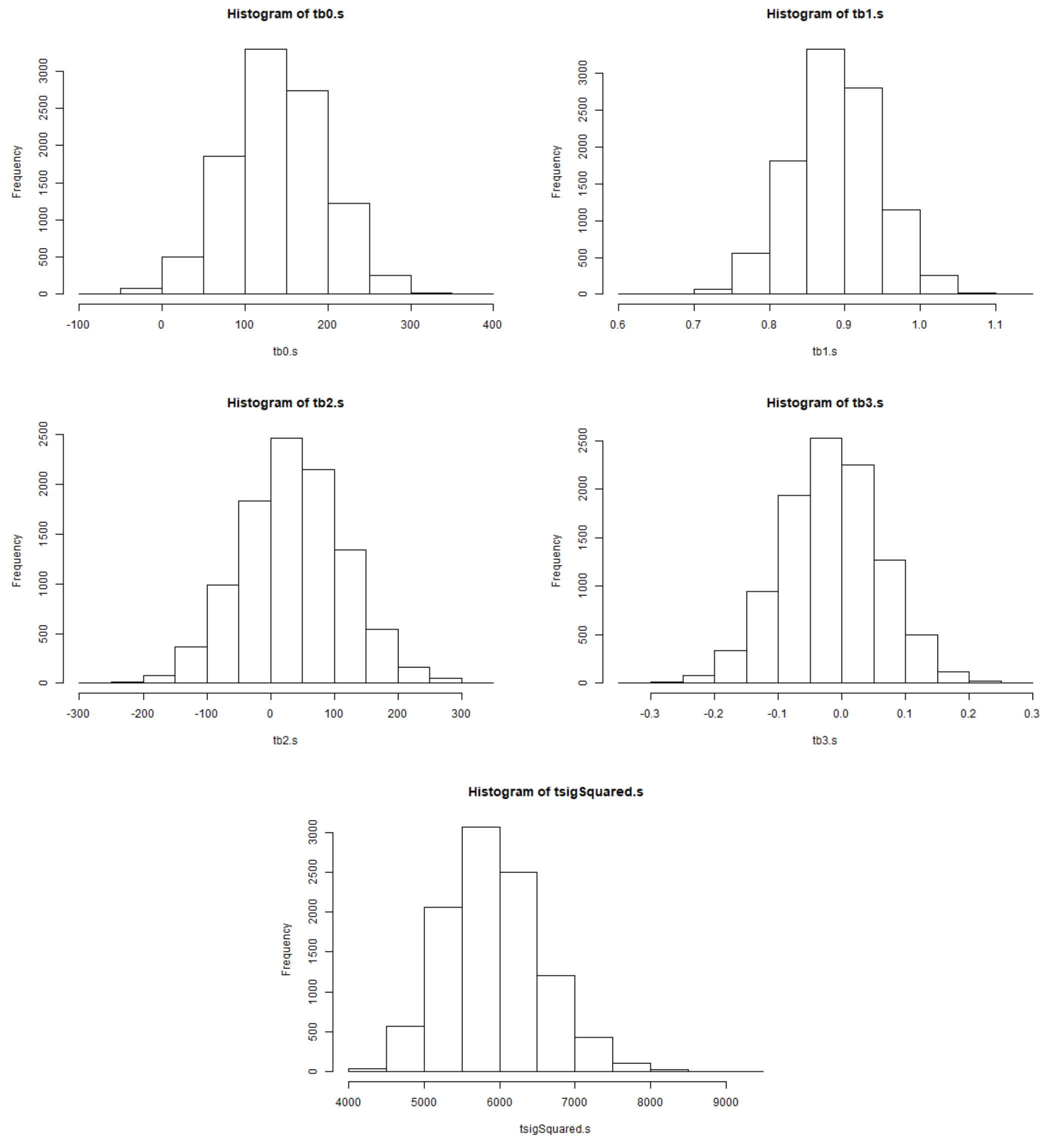


Figure A.11: Histograms of b_0, b_1, b_2, b_3, and σ^2 marginal posterior realizations for the posterior regression lines of the total PSAT and SAT scores.

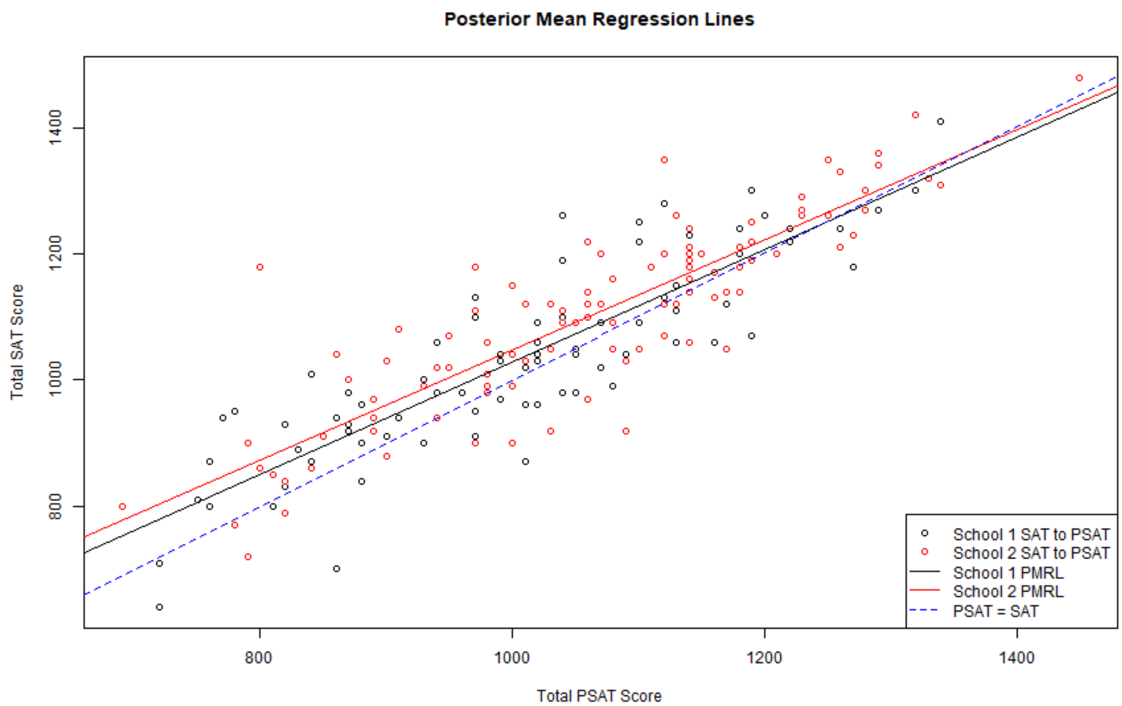
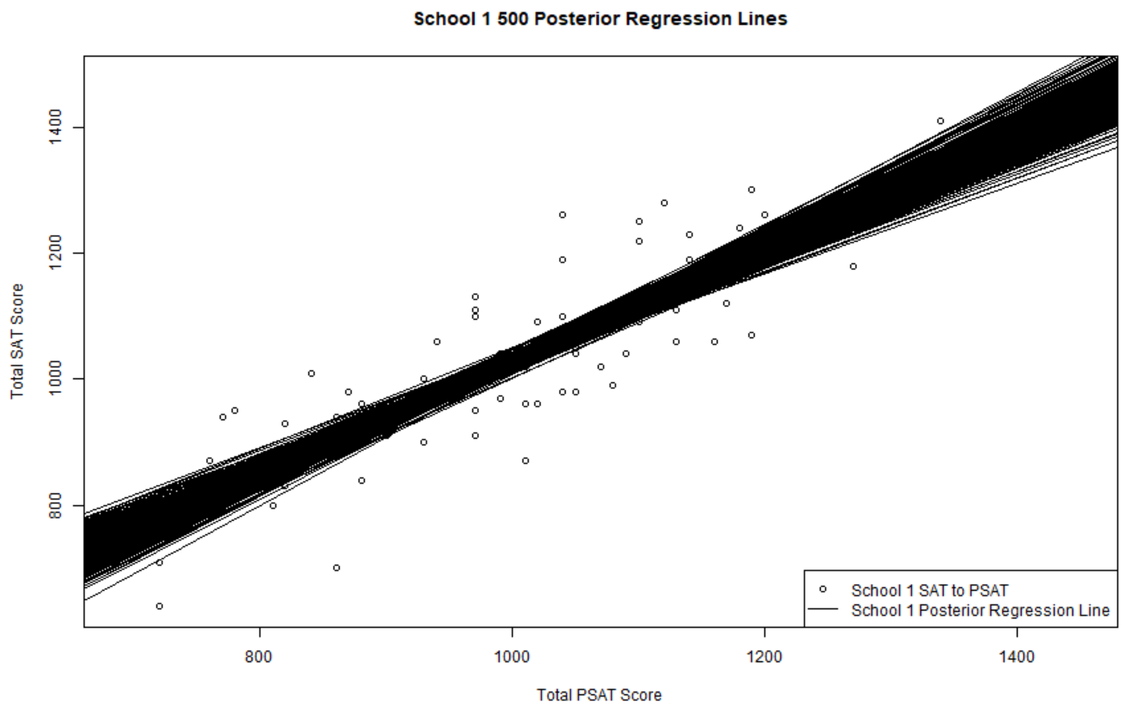


Figure A.12: Posterior mean regression lines of the total SAT score to the total PSAT score for both schools. School 1 total score data points are in black and School 2 total score data points are in red. The black and red lines are the posterior mean regression lines of School 1 and School 2 respectively. The blue, dashed line is the 45° line representing equality between total PSAT score and total SAT score.



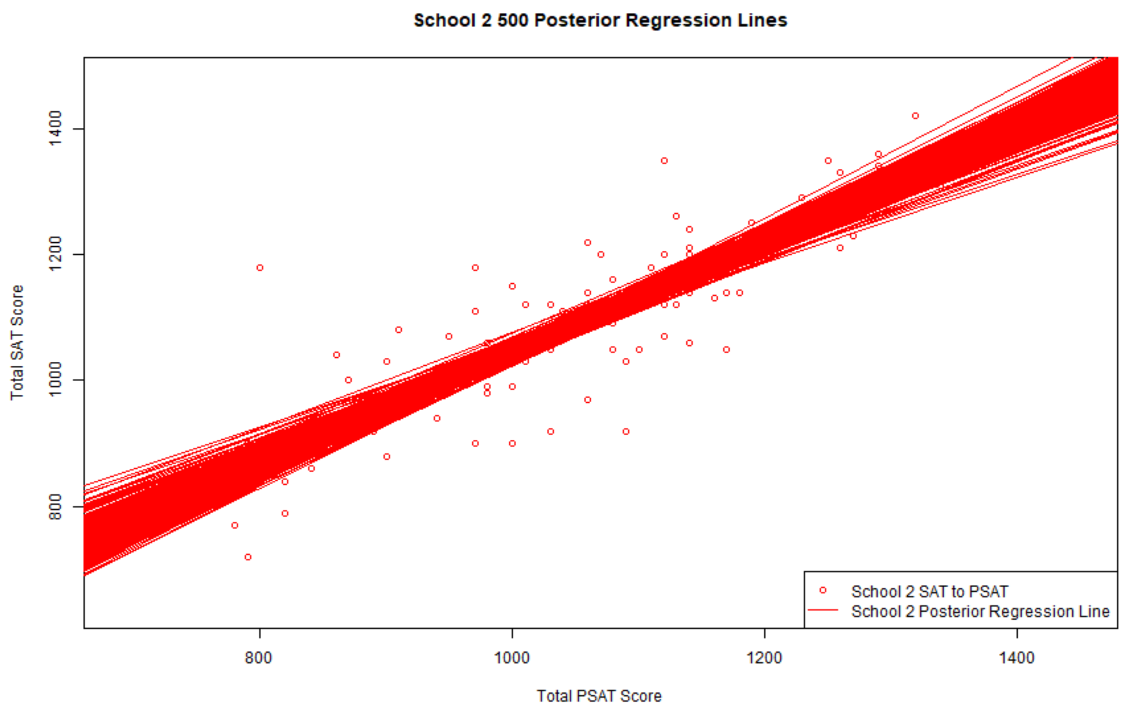


Figure A.13: The lines in these figures are 500 out of 10,000 of the posterior regression lines for both schools. School 1 and School 2 total score data points are in black and red respectively. Posterior regression lines for School 1 and School 2 are in black and red respectively.

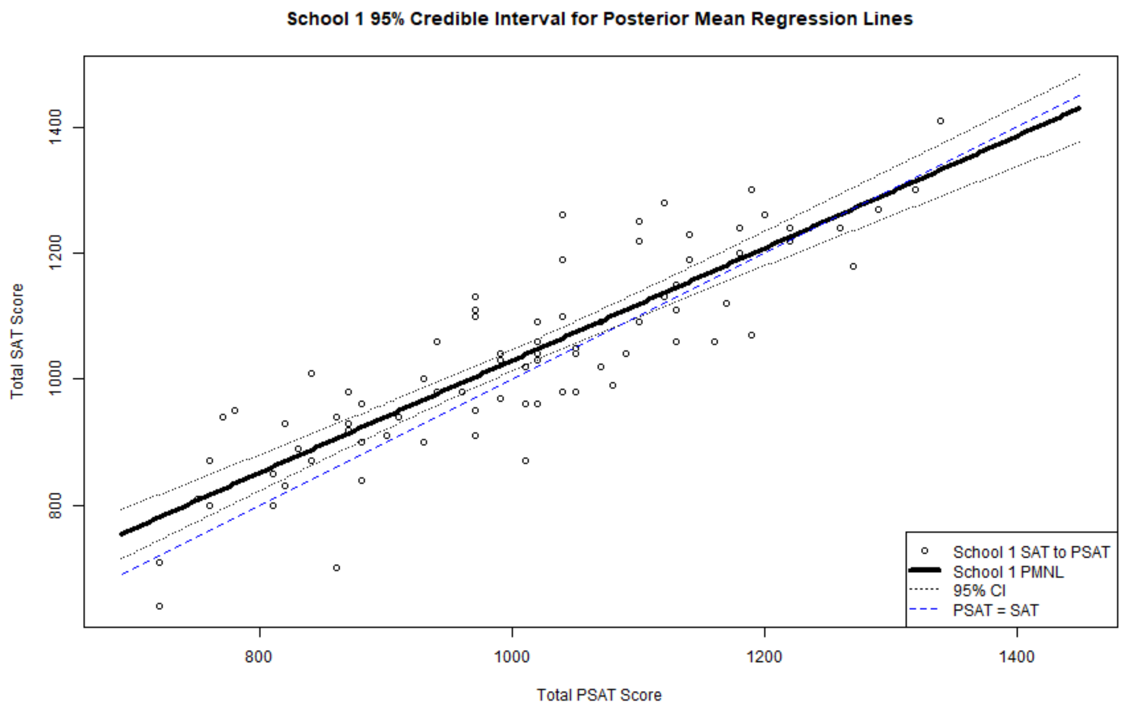
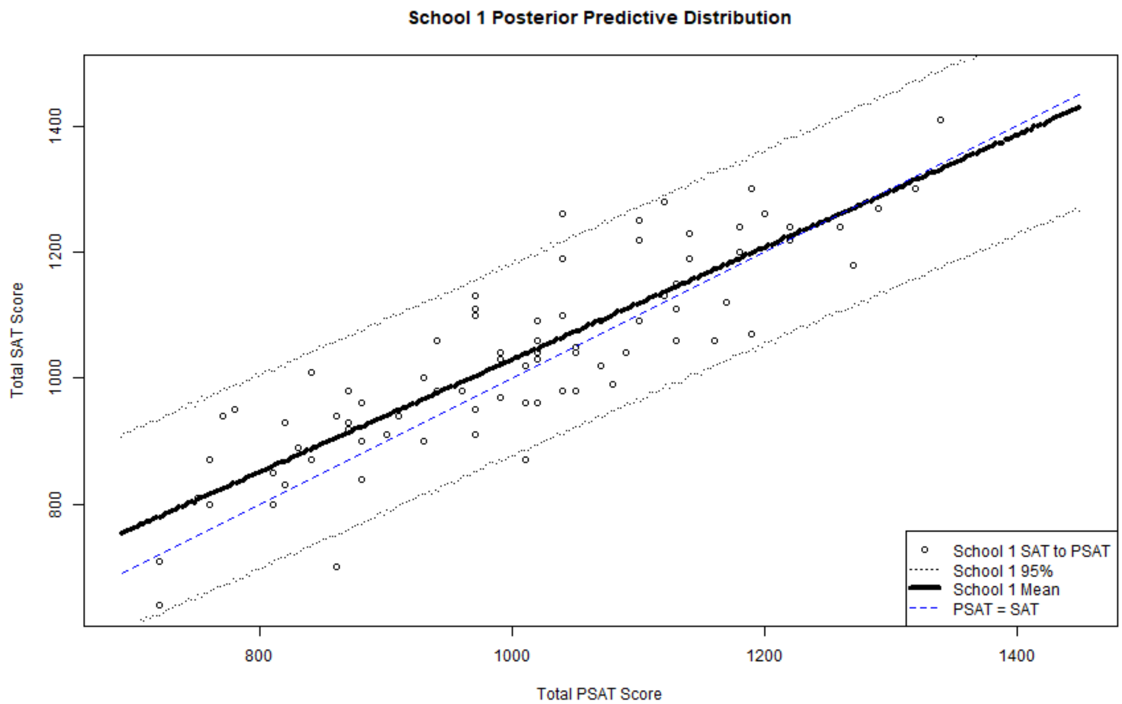


Figure A.14: School 1 and School 2 total score data points are in black and red respectively. The dotted curves represent the 95% credible interval for the true posterior mean regression line. The thick line is the mean posterior regression line.



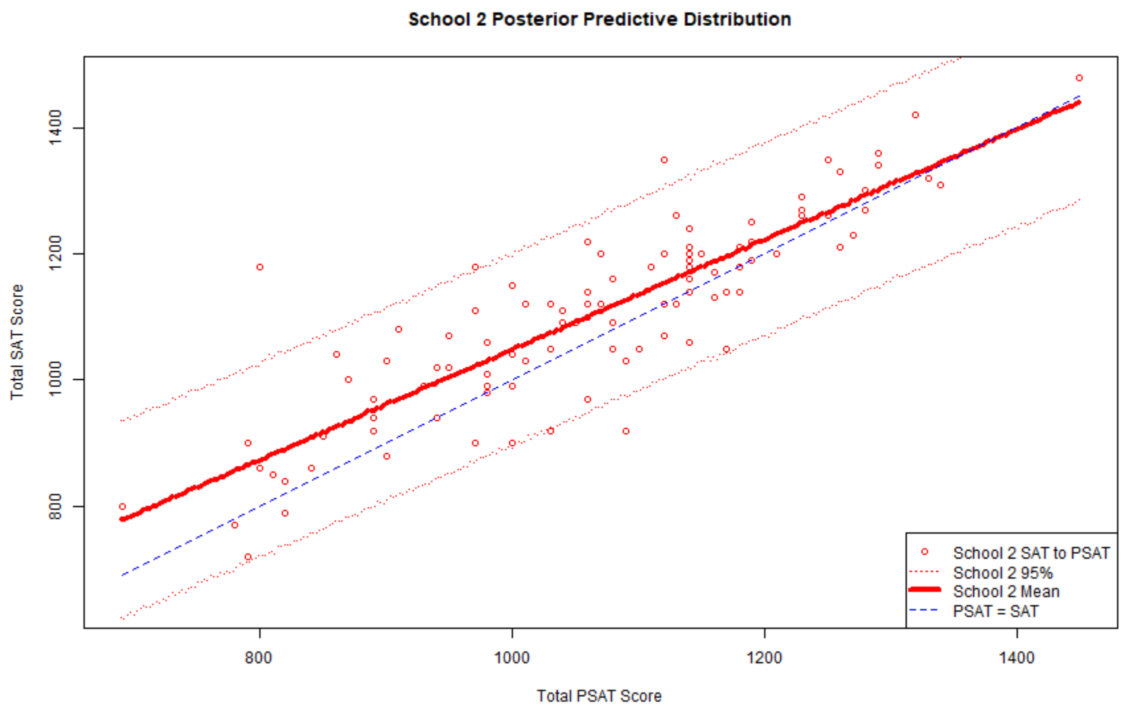


Figure A.15: School 1 and School 2 total score data points are in black and red respectively. The thick black and red lines represent the mean total score a student got on the SAT based on the PSAT for School 1 and School 2 respectively. The dotted black and red lines represent the 95% credible interval for a student's total SAT score based upon their total PSAT score. The blue, dashed line is the 45° line representing equality between total PSAT score and total SAT score.

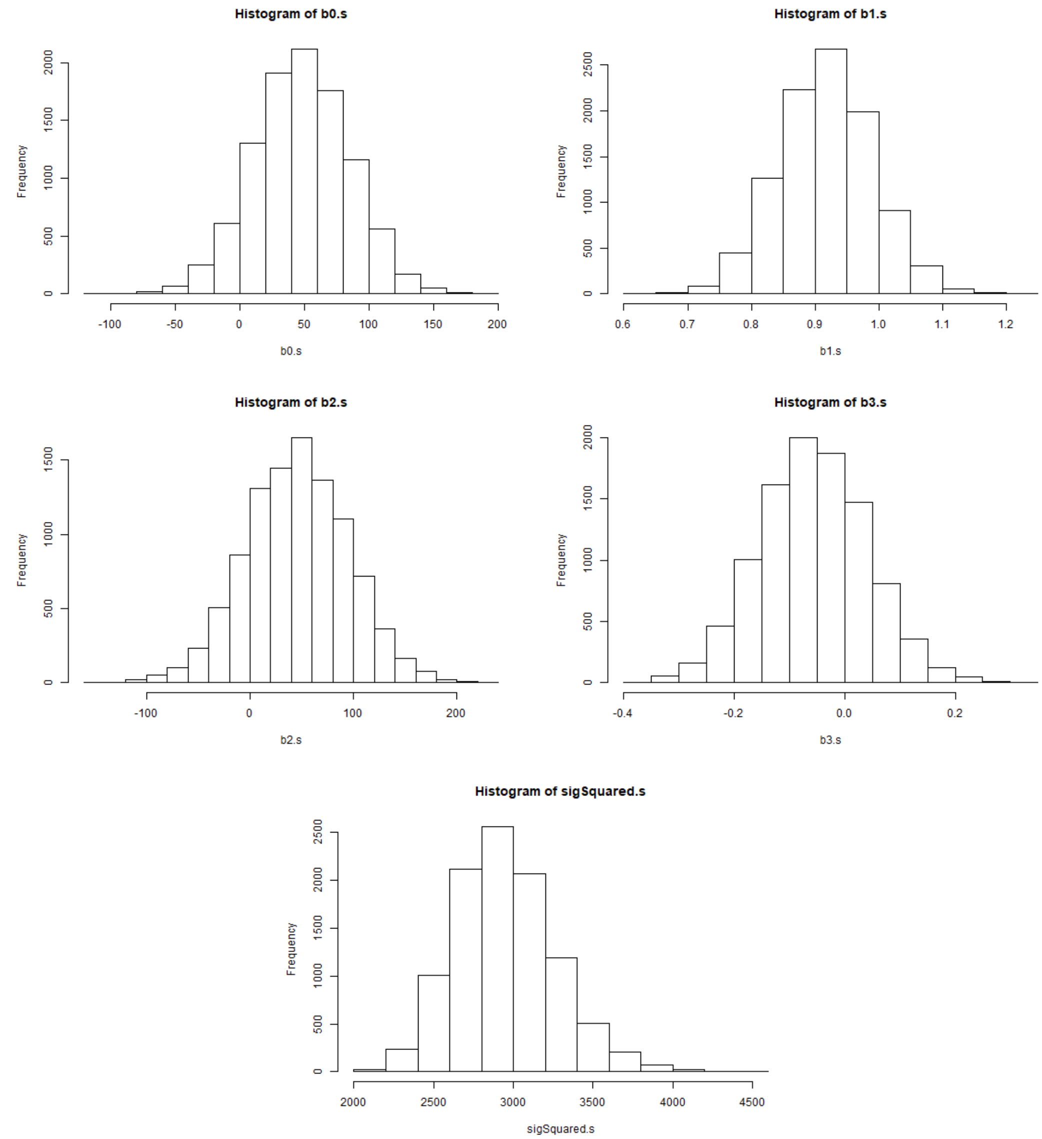


Figure A.16: Histograms of b_0, b_1, b_2, b_3, and σ^2 marginal posterior realizations for the posterior regression lines of the mathematics section scores of the PSAT and SAT.

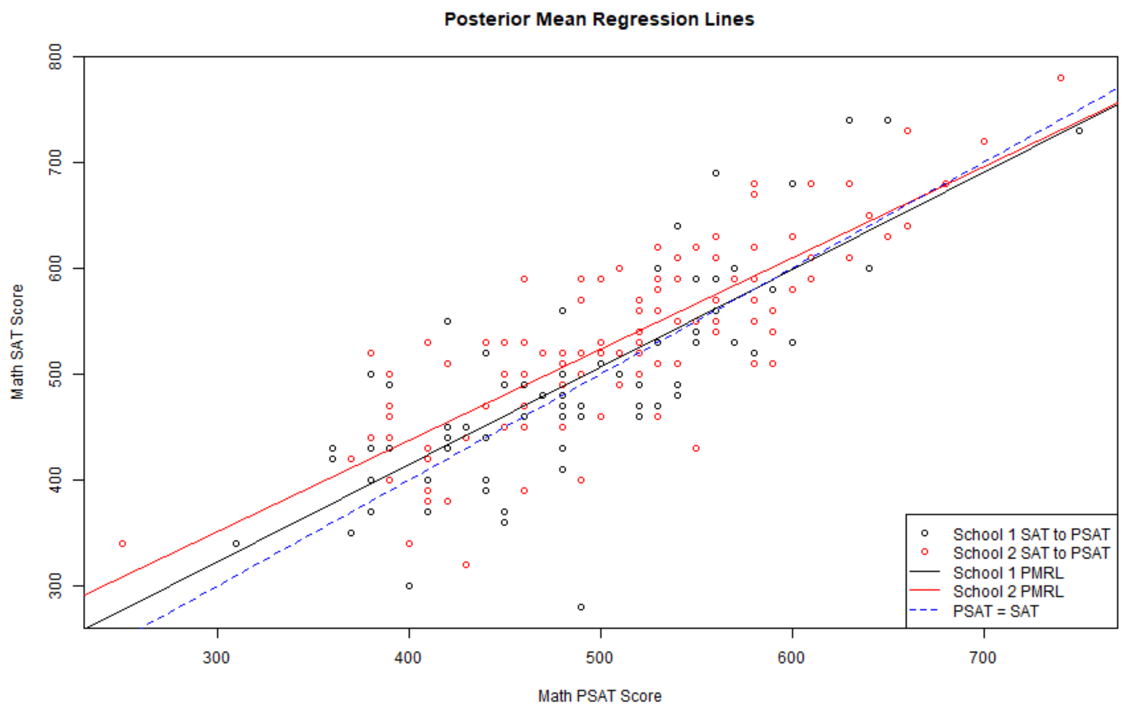


Figure A.17: Posterior Mean Regression Lines of the mathematics section PSAT score to the mathematics section SAT score for both schools. School 1 mathematics section score data points are in black and School 2 mathematics section data points are in red. The black and red lines are the posterior mean regression lines of School 1 and School 2 respectively. The blue, dashed line is the 45° line representing equality between mathematics section PSAT score and the mathematics section SAT score.

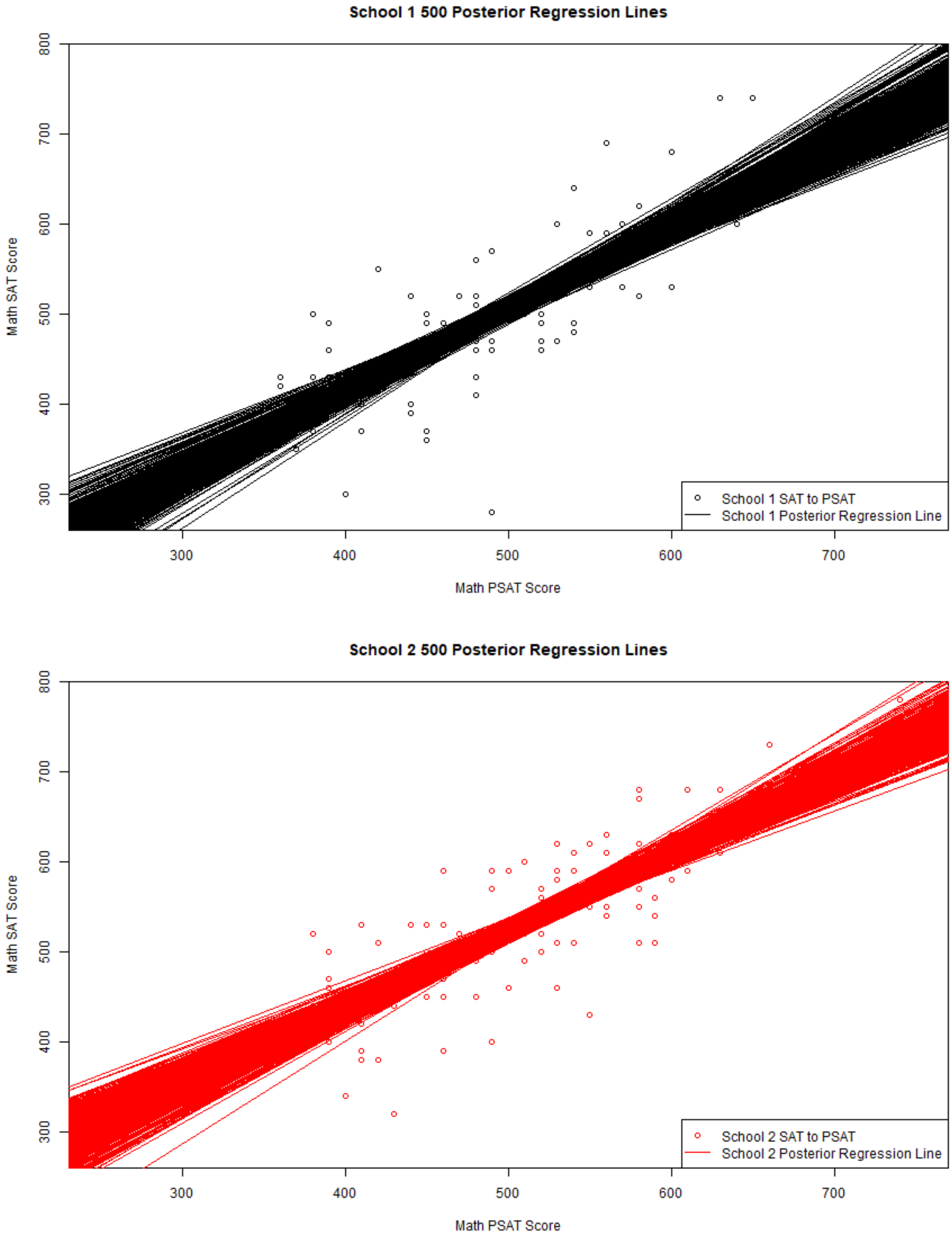
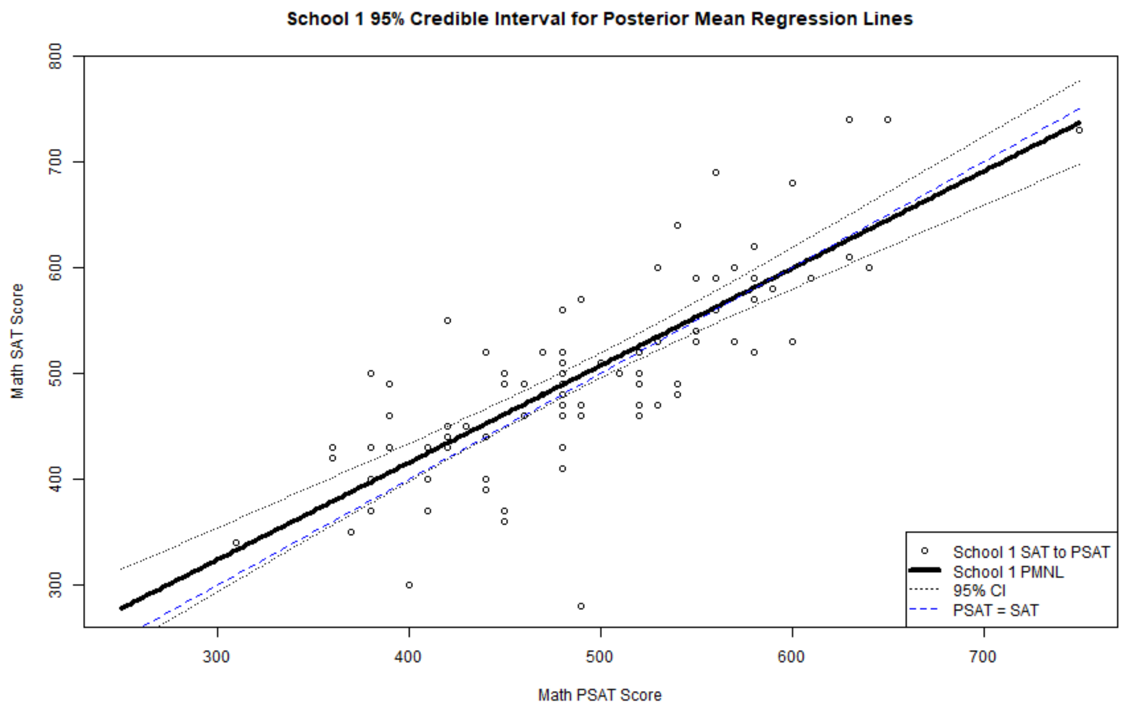


Figure A.18: These figures are 500 out of 10,000 of the posterior regression lines for the two schools. School 1 and School 2 mathematics score data points are in black and red respectively. Posterior regression lines for School 1 and School 2 are in black and red respectively.



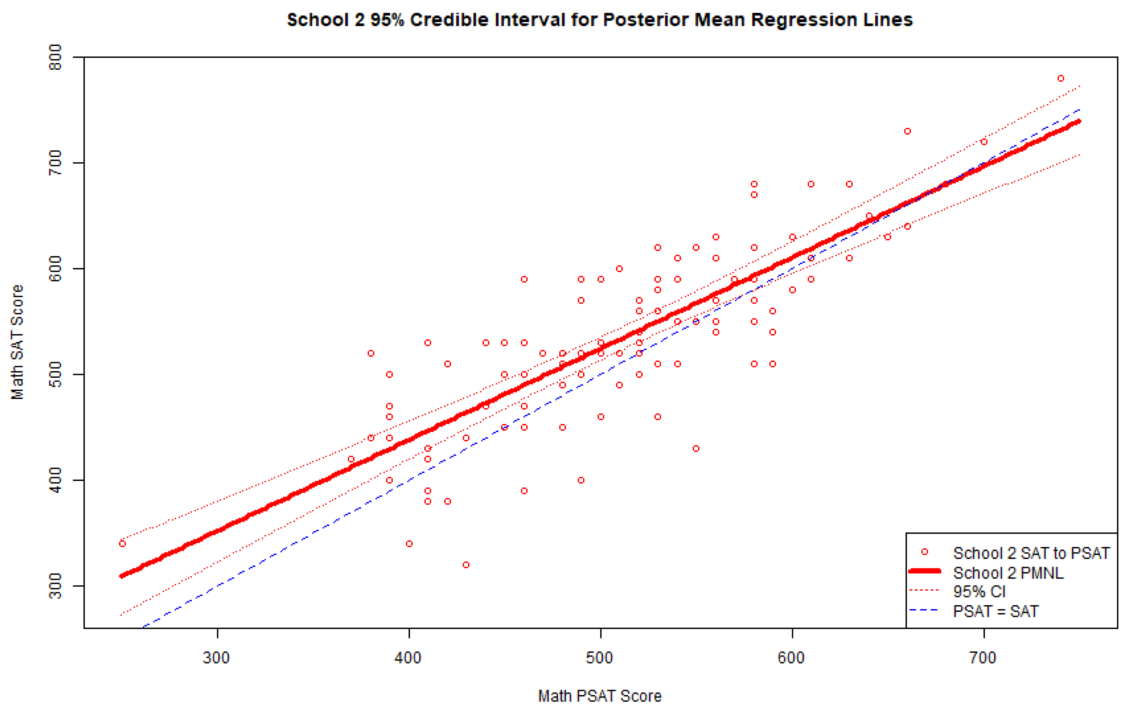
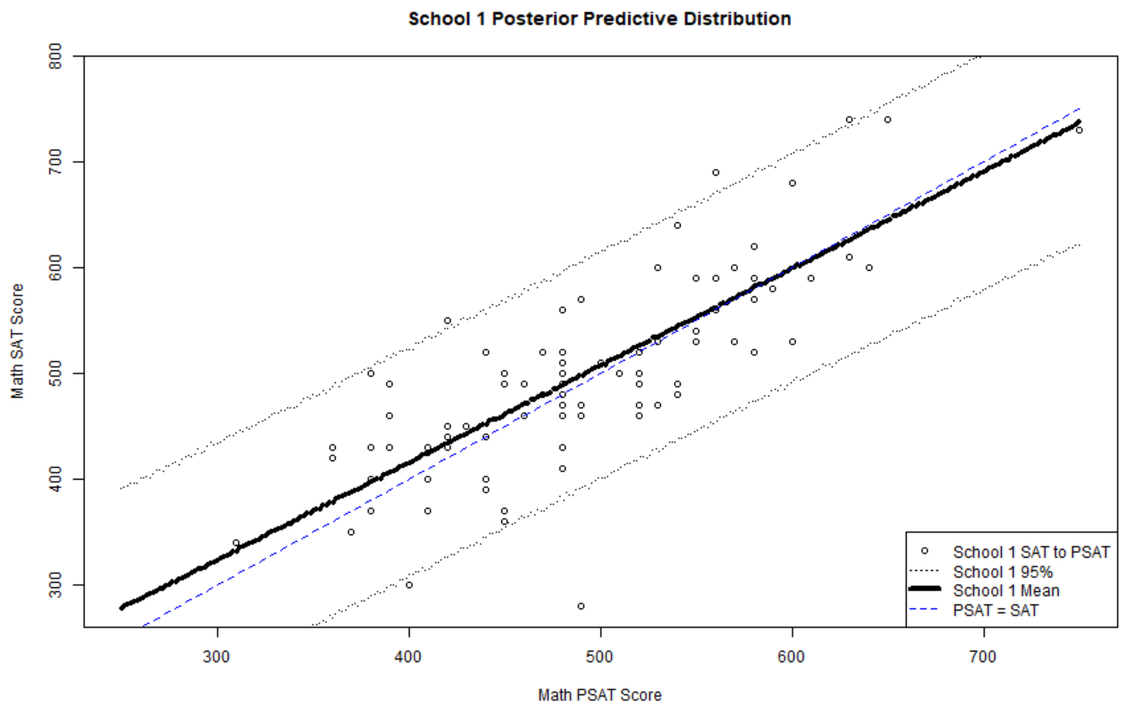


Figure A.19: School 1 and School 2 total score data points are in black and red respectively. The dotted curves represent the 95% credible interval for the true posterior mean regression line. The thick line is the mean posterior regression line. The blue, dashed line is the 45° line representing equality between mathematics PSAT score and the mathematics SAT score.



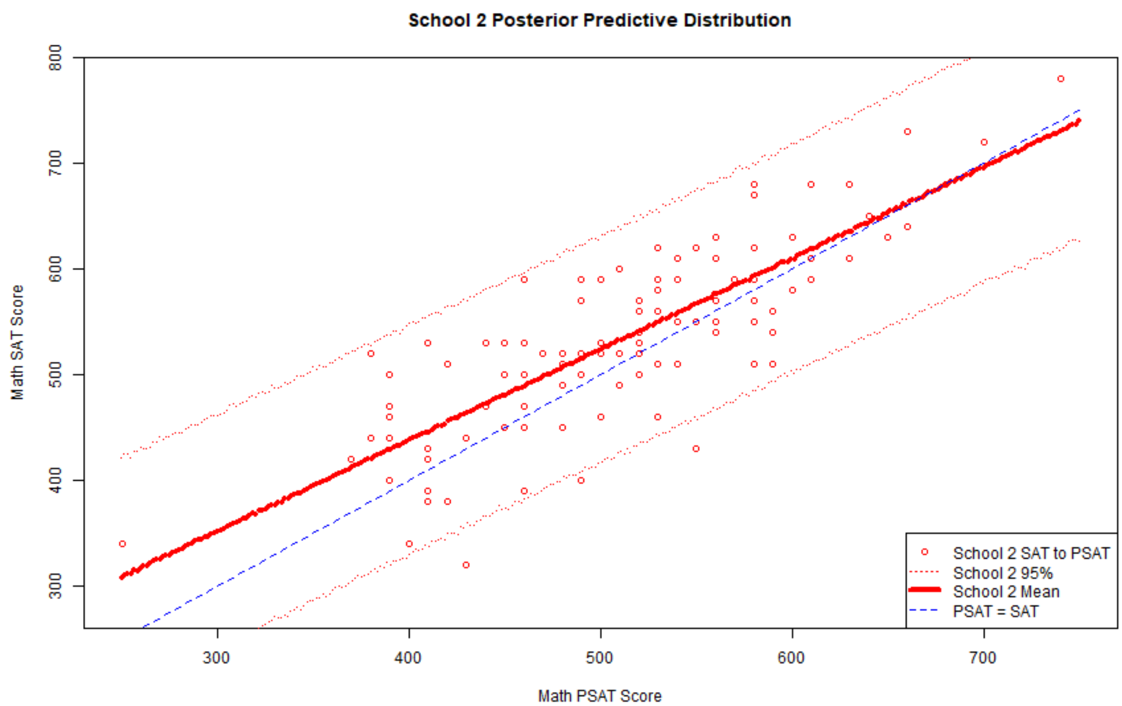


Figure A.20: The mathematics scores from School 1 are in black and the mathematics scores from School 2 are in red. Lines in black are from School 1 and lines in red are from School 2. The dotted curves are the 95% credible interval and the bold lines are the average score going from mathematics PSAT to mathematics SAT.

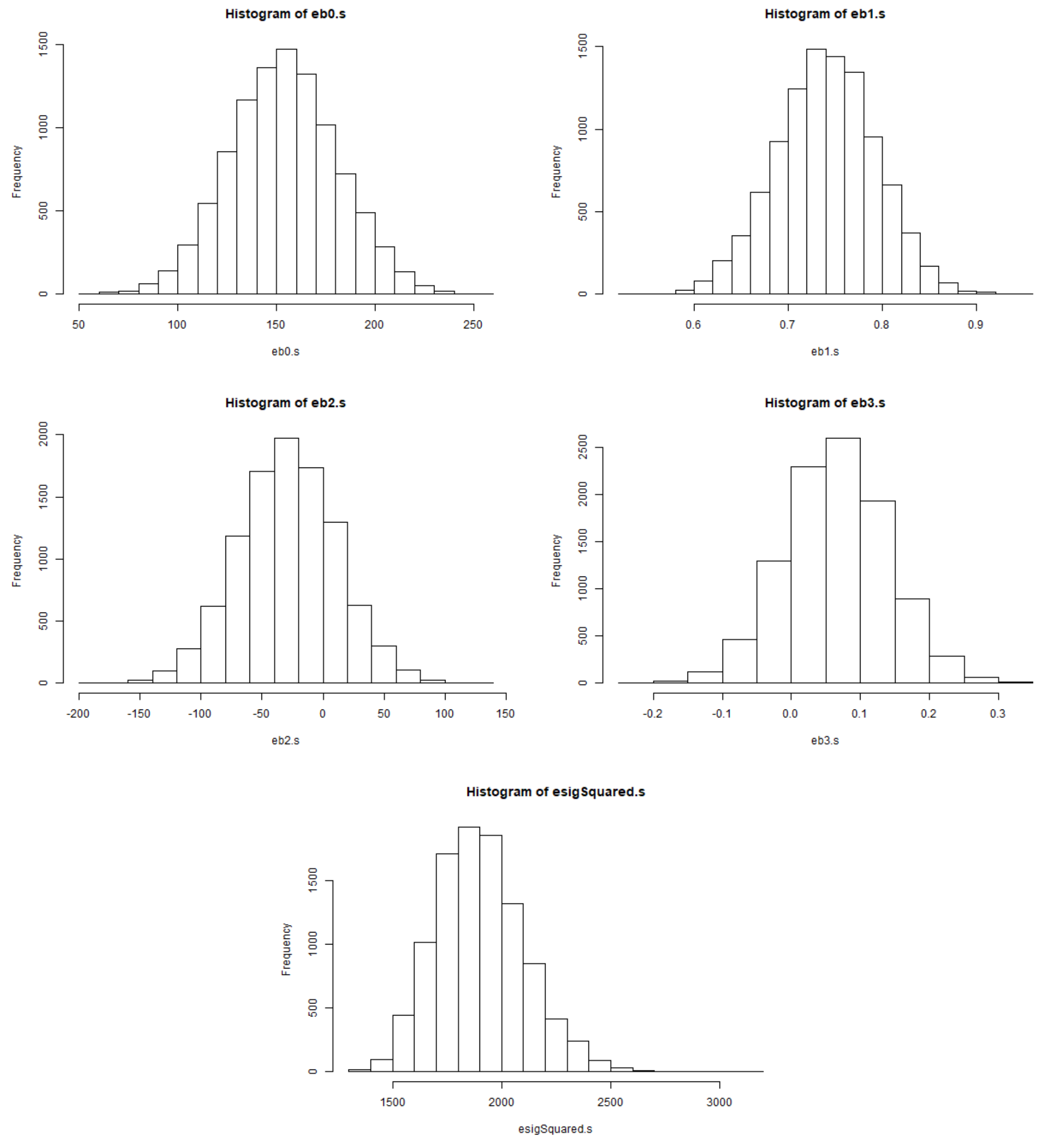


Figure A.21: Histograms of b_0, b_1, b_2, b_3, and σ^2 marginal posterior realizations for the posterior regression lines of the ERW scores of the PSAT and SAT.

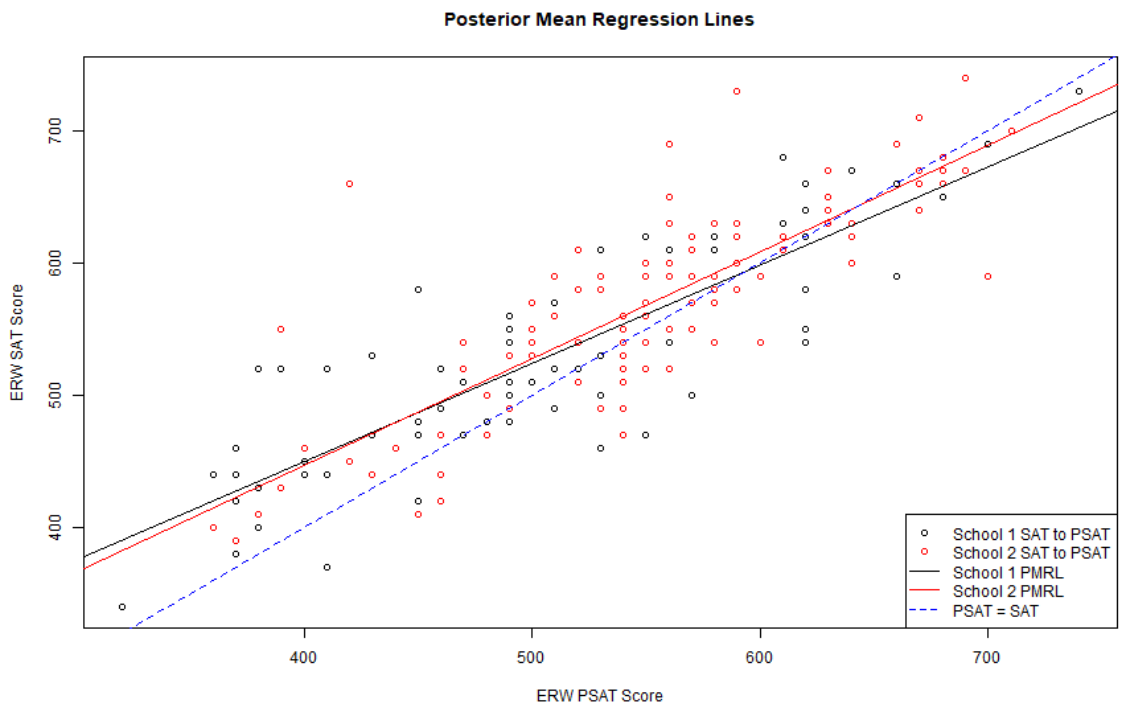
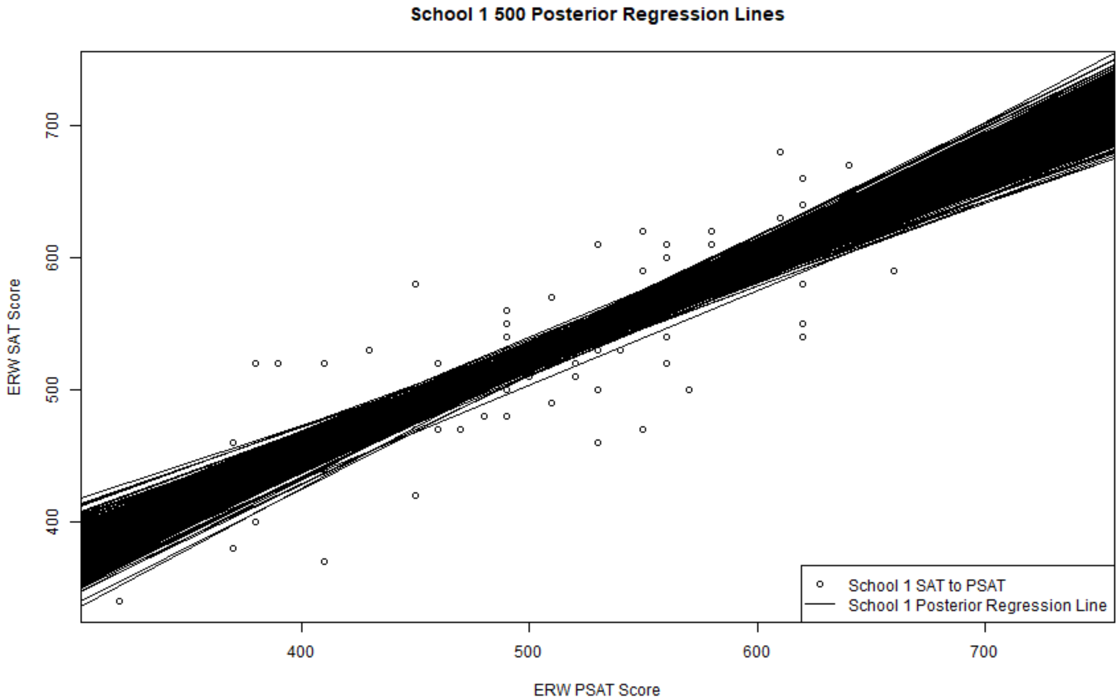


Figure A.22: Posterior Mean Regression Lines of the ERW PSAT score to the ERW SAT score for both schools. School 1 ERW score data points are in black and School 2 ERW data points are in red. The black and red lines are the posterior mean regression lines of School 1 and School 2 respectively. The blue, dashed line is the 45° line representing equality between ERW PSAT score and the ERW SAT score.



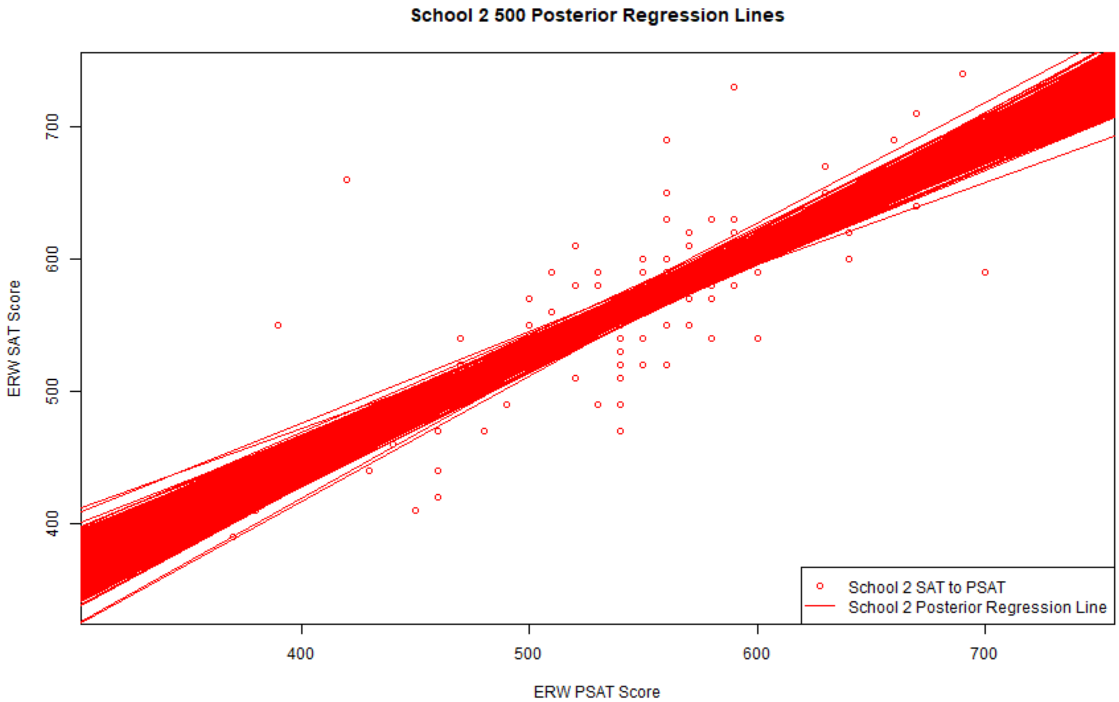
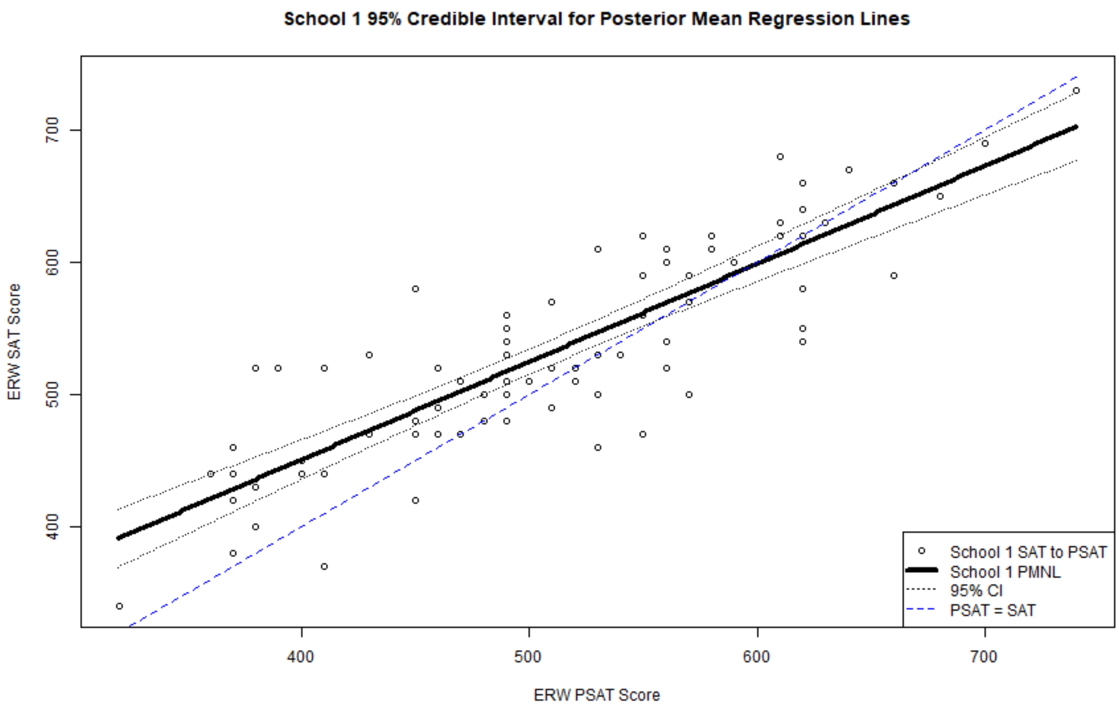


Figure A.23: These figures are 500 out of 10,000 of the posterior regression lines for the two schools. School 1 and School 2 ERW score data points are in black and red respectively. Posterior regression lines for School 1 and School 2 are in black and red respectively.



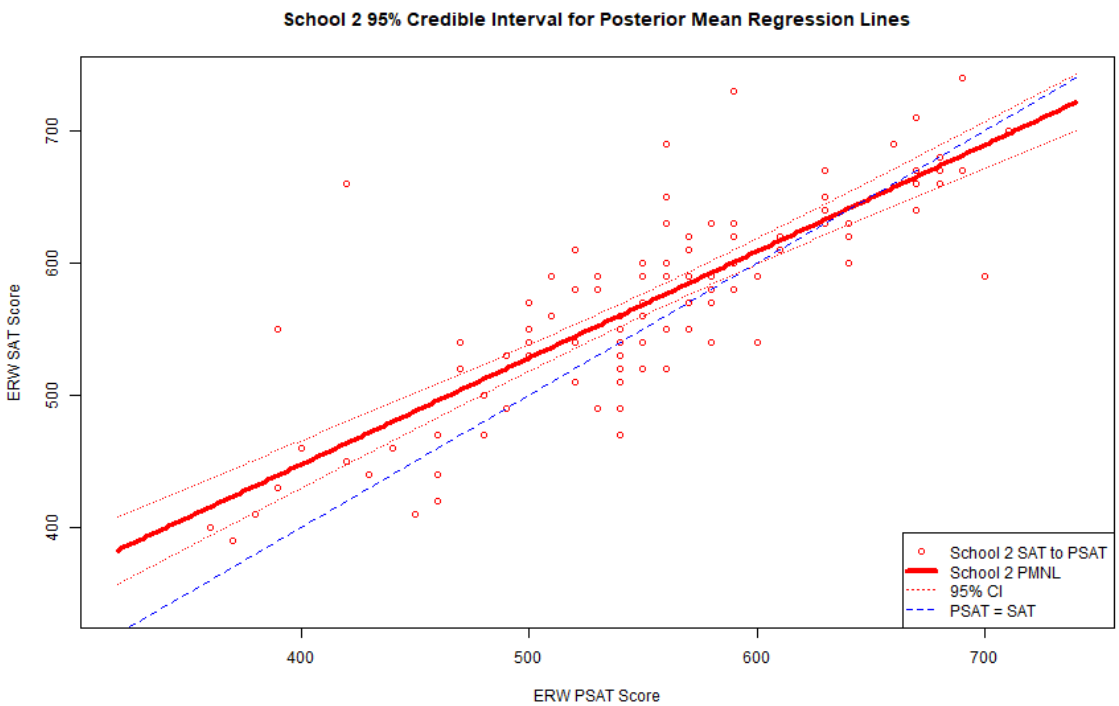
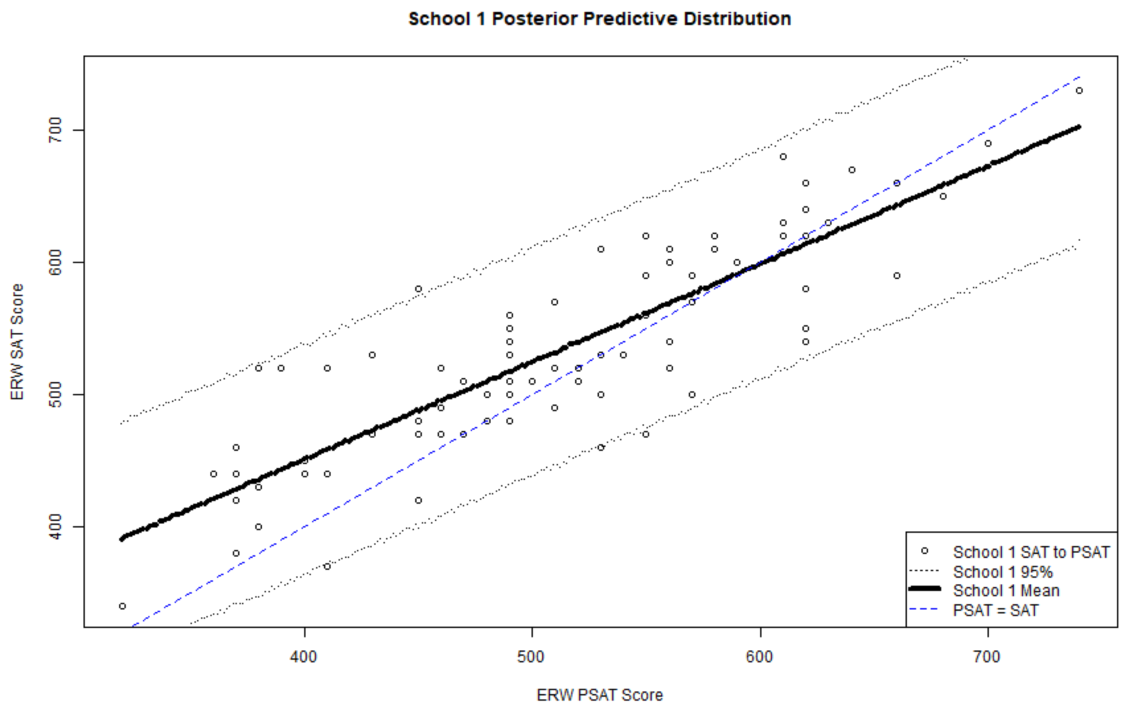


Figure A.24: School 1 and School 2 total score data points are in black and red respectively. The dotted curves represent the 95% credible interval for the true posterior mean regression line. The thick line is the mean posterior regression line. The blue, dashed line is the 45° line representing equality between ERW PSAT score and the ERW SAT score.



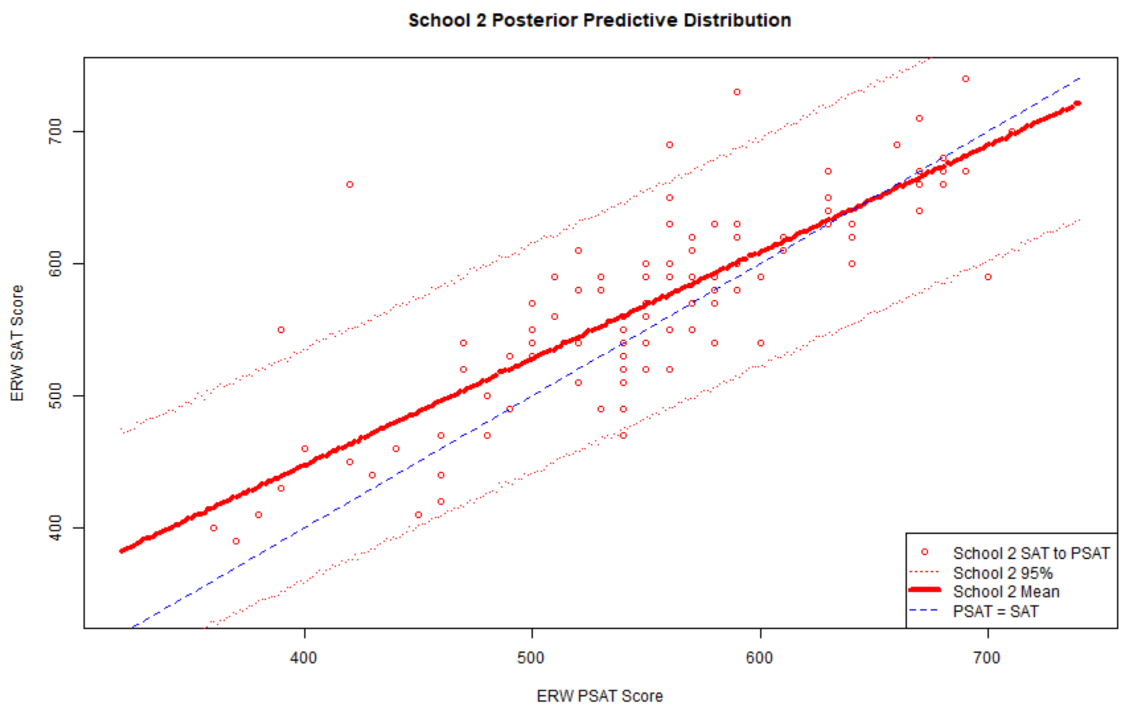


Figure A.25: The ERW scores from School 1 are in black and the ERW scores from School 2 are in red. Lines in black are from School 1 and lines in red are from School 2. The dotted curves are the 95% credible interval and the bold lines are the average score going from ERW PSAT to ERW SAT.

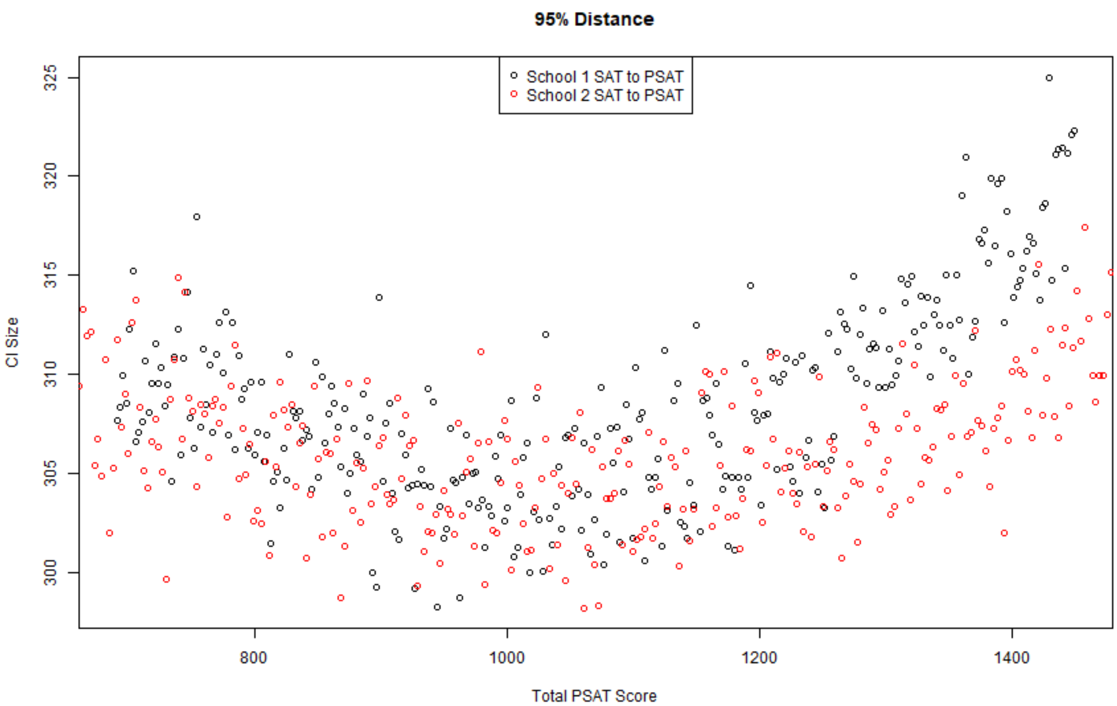


Figure A.26: The vertical distance between the 97.5% quantile and the 2.5% quantile for total scores. Distances gathered from School 1 are represented by back circles and distances gathered from School 2 are represented by red circles.

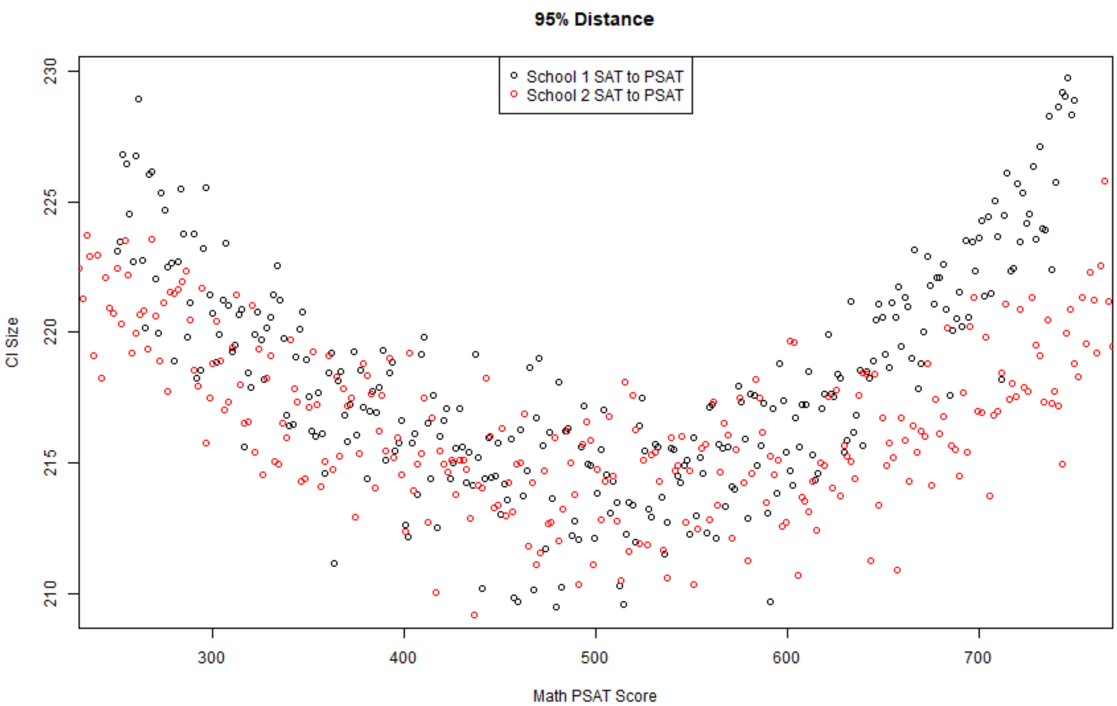


Figure A.27: The vertical distance between the 97.5% quantile and the 2.5% quantile for scores on the math section. Distances gathered from School 1 are represented by back circles and distances gathered from School 2 are represented by red circles.

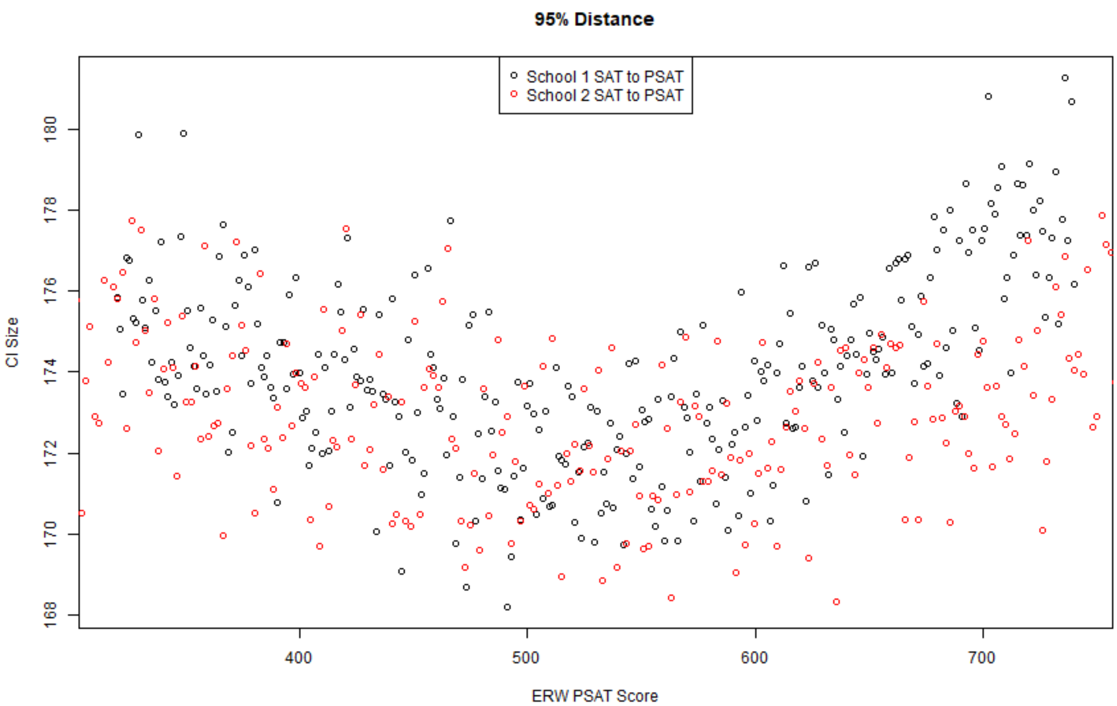


Figure A.28: The vertical distance between the 97.5% quantile and the 2.5% quantile for scores on the ERW section. Distances gathered from School 1 are represented by back circles and distances gathered from School 2 are represented by red circles.

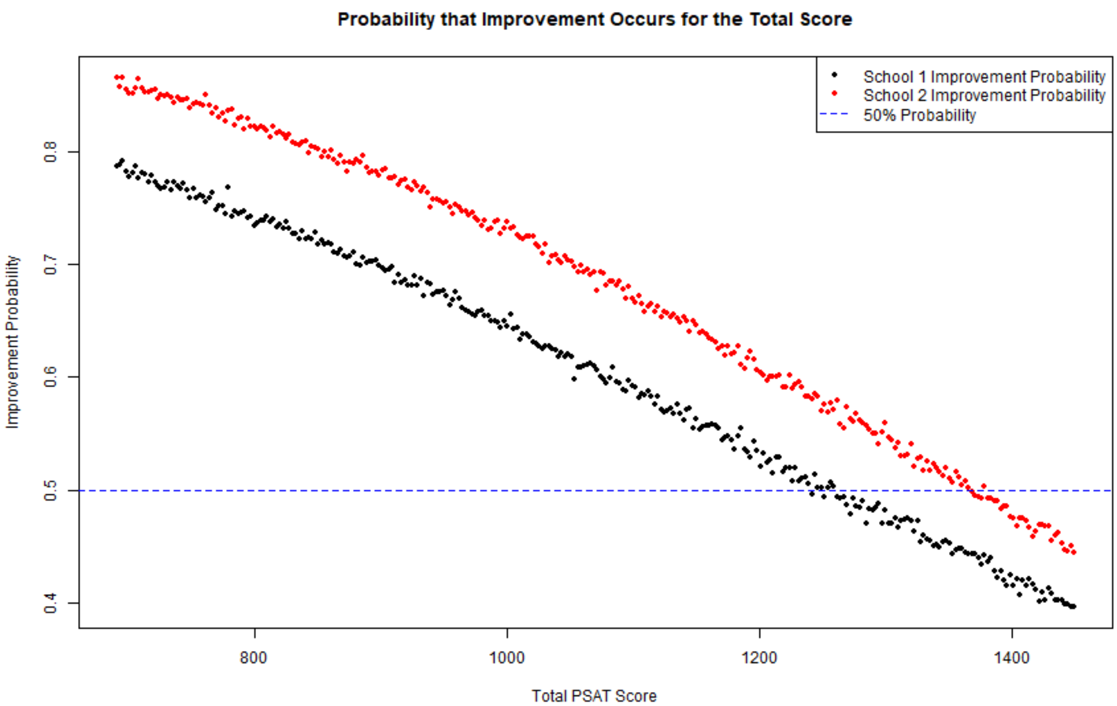


Figure A.29: The probability that a student will get a total score on their SAT that is better than their PSAT total score, as separated by School 1 (black), and School 2 (red).

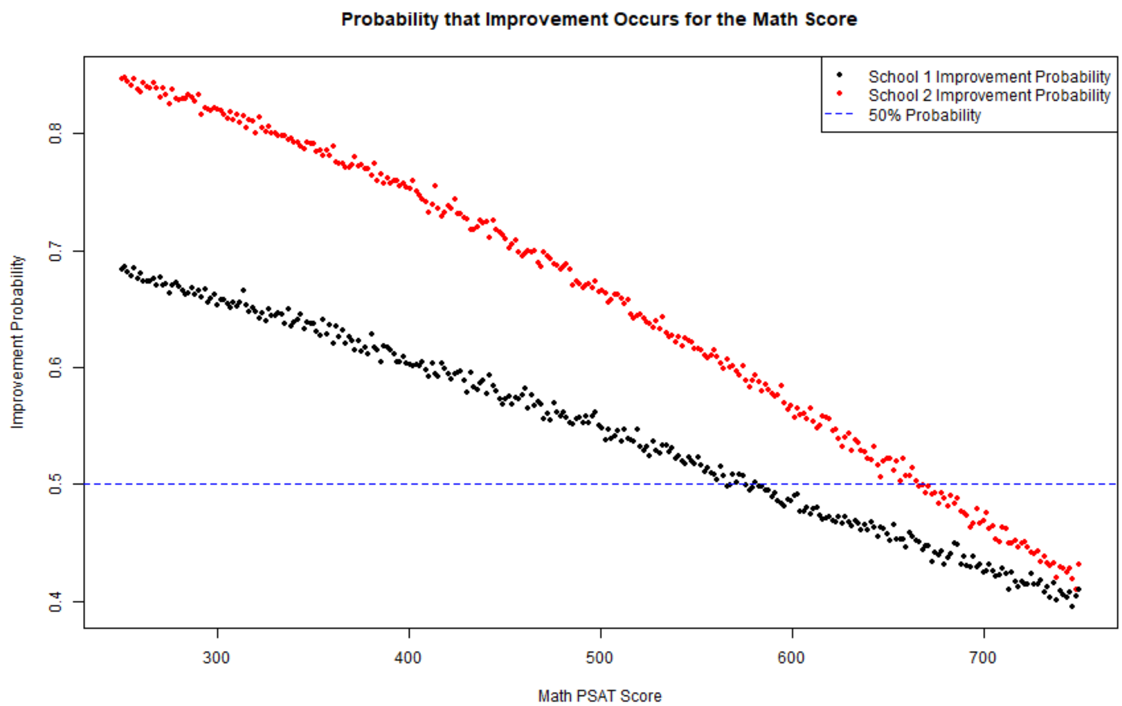


Figure A.30: The probability that a student will get a math score on their SAT that is better than their PSAT math score, as separated by School 1 (black), and School 2 (red).

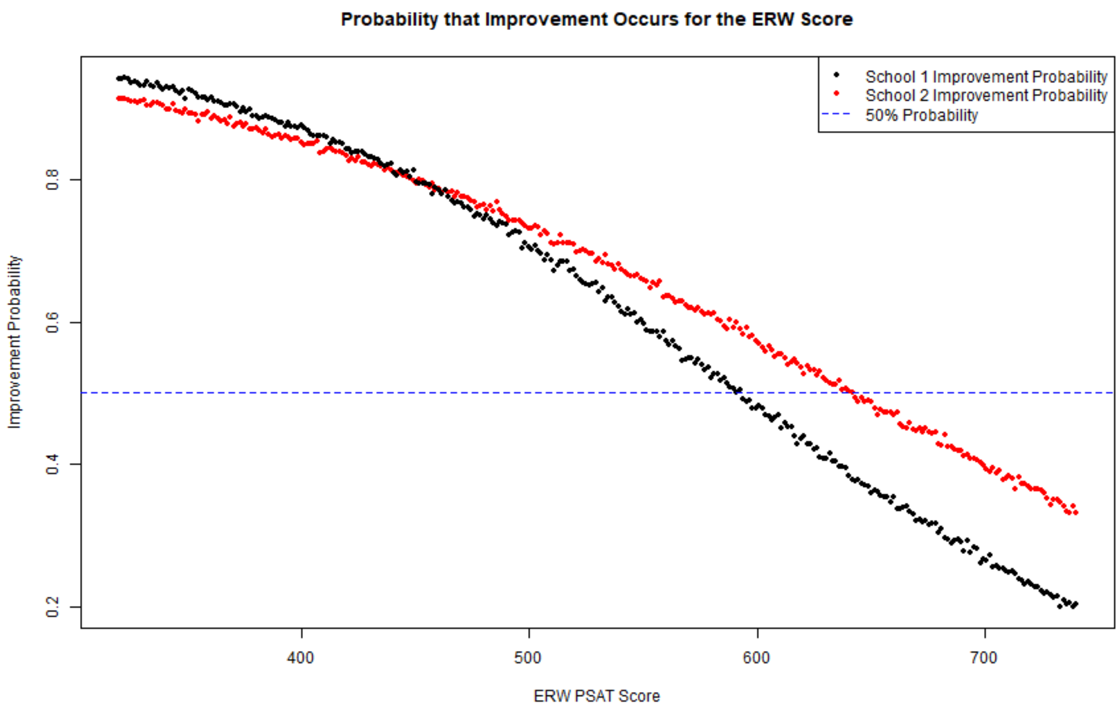


Figure A.31: The probability that a student will get an ERW score on their SAT that is better than their PSAT ERW score, as separated by School 1 (black), and School 2 (red).