Based on feedback from my peer reviewer, I made adjustments to move sections from the introduction that explained my decision to select certain questions in my model to the methods section. This structural choice also made sense to me because the methods section now includes a more complete description of my model and the reasoning behind it. I also included more analysis in my appendix and explanation for validating the models in terms of normality and constant variance. This section is added to the methods section. My peer reviewer also recommended I use Spearman correlation coefficient over Kendall's Tau, which I implemented in my results and discussion. In addition, I included a discussion of potential follow up research to conclude my paper.

Based on feedback from the instructor, I removed content that was wordy and unnecessary. This made the paper more clear, condensed, and straightforward.

Sensitivity Analysis of School Security Measures

By Emily MacQuarrie

Introduction:

Implementing security measures and enforcing standards is one method for decreasing the number of schools affected by school shootings and other safety-related concerns. Finn and Servoss's paper intends to address the relationship between school security and racial and gender demographics as well as to understand disparities between how schools implement security measures. They argue that larger schools, schools with a higher proportion of black students, schools with higher percentages of students on free lunch, and schools with higher percentages of students suspended all have higher security measures (Finn, Servoss 11). However, these claims lose legitimacy if the model they produce is not valid.

Finn and Servoss construct a Rasch model to understand how security compares across schools who participated in a school administrator survey. Data was based on a sample of K-12 U.S. school surveys. This questionnaire was part of the Education Longitudinal Study of 2002 sponsored by the Department of Education and the National Center for Education Statistics. Finn and Servoss selected 7 questions which they believed to be the best predictors for school security. These questions included information about whether schools require students to pass through metal detectors each day (q38c), perform one or more random metal detector checks on students (q38d), require drug testing for any students (q38h), perform one or more random sweeps for contraband but not including dog sniffs (q38g), use one or more security cameras to monitor the school (q38n), regularly use paid law enforcement or security services at school at any time during school hours (q40a), and use one or more random dog sniffs to check for drugs (q38f).

This paper aims to verify the validity of Finn and Servoss's model (Model 1) by reconstructing it and comparing the results against those of a new, justifiable model (Model 2). I conduct a sensitivity analysis of Finn and Servoss's study to understand the reliability of the model in question. A drastically different result in my own model would indicate that Finn and Servoss's methods are not reliable, and an identical result would confirm the reliability of their model.

Methods:

In order to test the reliability of the study conducted by Finn and Servoss, the model described in their paper was replicated. First, schools with missing responses were removed from the analysis to prepare the data for modeling. Missing values were omitted as opposed to imputed because only 19 schools out of 656 were removed, which is a relatively small proportion. Moreover, the data is assumed to be missing at random. In other words, for each school, question responses that were missing showed no pattern in missingness. A Rasch model with binomial logistic regression was fit to this dataset. This type of model is used over others because it accounts for questions that are more or less likely to have a 'yes' response in the prediction; not all questions are weighted equally. The 7 survey questions as well as the school IDs were used as predictors in the model, and the outcome variable was the school responses.

Model 2 also assumes data is missing at random for the same reasons mentioned above. Missing values were again omitted, and 21 out of the 656 entries were removed. The same

Rasch model was fit with a separate subset of 7 security questions. These questions asked schools about varying levels of security they maintained: control access to school buildings during school hours (q38a), close the campus for most students during lunch (q38e), perform one or more random sweeps for contraband but not including dog sniffs (q38g), require drug testing for any students (q38h), require faculty and staff to wear badges or picture IDs (q38m), use paid law enforcement security services at selected school activities (q40c), and use paid law enforcement security services when school and related activities were not occurring (q40d).

These selected questions were based on research conducted by the National Longitudinal Study of Adolescent to Adult Health. Their research shows that visible security measures, such as metal detectors and cameras, caused students to feel unsafe (Perumean-Chaney and Sutton). Security should not only protect students from physical harm, but it should also ensure students are mentally and emotionally at ease. For this reason, I chose questions where security measures were less visible or intrusive.

Both models were tested for goodness-of-fit using Pearson's Chi-Square Test. To conduct the sensitivity analysis, the resulting ranks of the schools from Model 1 were graphed against the ranks of the schools from Model 2. The ranks of the schools are determined by a security index. The ranks are produced by ordering the predicted school coefficients from the model. The correlation in these results determines model validity. The Spearman correlation coefficient was calculated to quantify this correlation.

I also tested slight variations of these two models. Both were fit with random effects instead of fixed effects, and the intercepts of these random effects models acted as the security indices. To compare fixed effects against random effects, the resulting school indices from both models were plotted against each other.

The random effects version of Model 2 was then evaluated with bootstrapping methods. 1000 repetitions of this model were run, and the variance of residuals and normality of the results were observed for a randomized subset of three schools. Variance was tested by taking the standard deviation of the difference between the bootstrapped indices and the actual index. Normality was tested by graphing the bootstrapped indices in a histogram. Variance and normality were evaluated because this is an assumption that was made for the Rasch model in the Finn and Servoss study. Testing these criteria on Model 2 further justifies the comparison of the two models and also verifies fitting a Rasch binomial logistic regression to the data.

Results:

Model 1 has a residual deviance of 3280.5 on 3816. Model 2 has a residual deviance of 3766.1 on 3804 degrees of freedom. The Chi-square statistic with a 95% confidence interval for the Model 1 is 3961, and the Model 2 evaluation statistic is 3949. Both models pass the Pearson's Chi-Square test for independence, therefore, we can say that the models are both statistically significant.

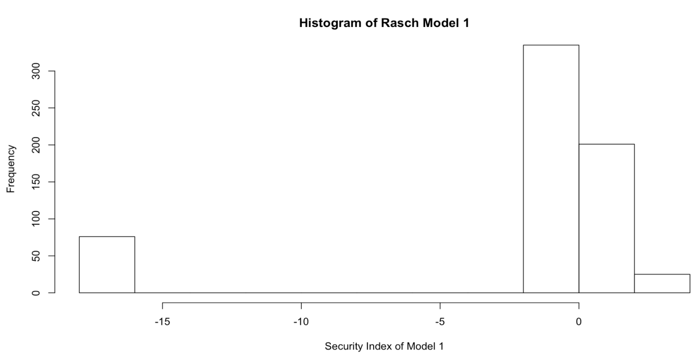
The distribution of the security indices are significantly different when comparing those of Model 1 to Model 2. The distribution of security indices from Model 1 appear to differ significantly from that of Model 2, which can be seen in Figures 1 and 2 respectively.

After plotting the ranks of the first model against the ranks of the second model, there appears to be a general positive trend, but variance is high. The Spearman correlation coefficient resulted in a value of 0.557. Because the correlation coefficient is not close to 1, the school security ranks of Model 1 and the school security ranks of Model 2 are not strongly correlated.

After plotting the indices of the random effects versus the indices of the fixed effects models, it appears that the random effects version of both Model 1 and Model 2 is a better representation of the data. In the fixed effects version, there is a wide dispersion of indices, which does not make sense in the context of school security. The random effects indices are closer together in value and show less disparity. This makes more sense because there does not typically exist a small number of schools that are drastically more or less secure than most schools. Thus, I chose to evaluate the random effects model over the fixed effects model.

Results of the random effects version of Model 2 after bootstrapping appeared to have normal distribution for all three selected schools. Additionally, the standard deviation of residuals was evaluated and yielded a similar value amongst all three schools. Thus, we can conclude that the fitted data satisfied normality and variance conditions.

Figure 1 (left) : This is a histogram of the security indices for model 1. There is a large distribution that is not close to Normal. Figure 2 (right) : This is a histogram of the security indices for model 2. Indices are highly frequent around 0, but a few are distributed around -20 and 20.



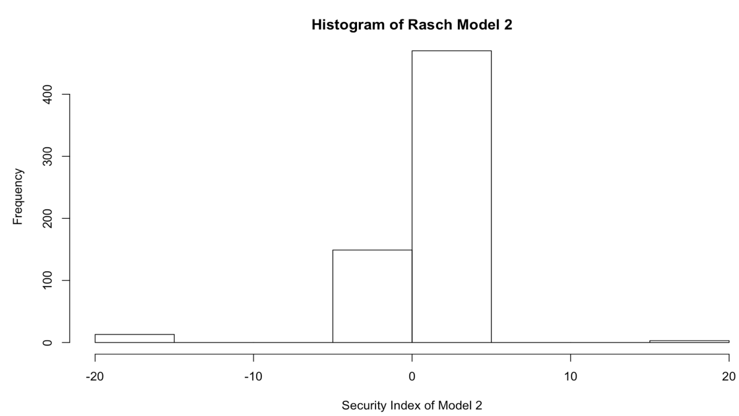
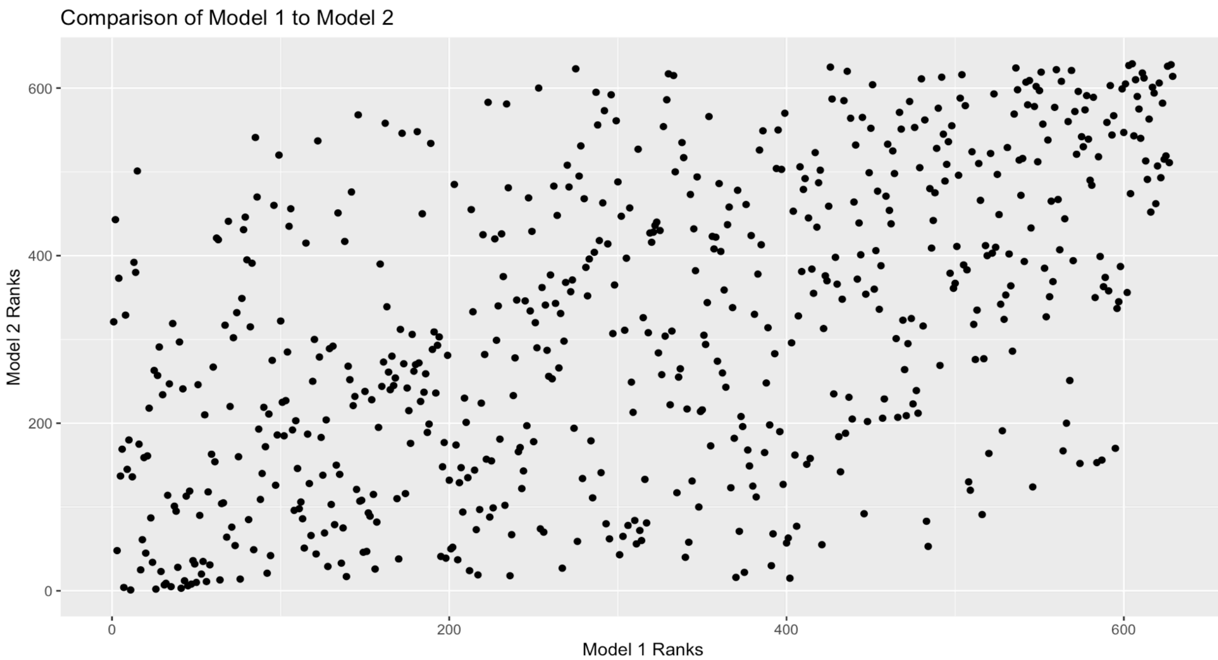


Figure 3: The scatterplot graphs the ranks of the first model on the x-axis, and the ranks of the second model on the y-axis. There appears to be a positive trend in the data with a high amount of variance.



Discussion:

By comparing the school ranks of Model 1 to the school ranks of Model 2, we can assess the validity of Finn and Servoss's study. A linear trend with high correlation would suggest that the questions chosen for the model are independent of the school security ranking. This finding would reinforce Finn and Servoss's work.

However, because the model constructed in this study yielded different results from Finn and Servoss's model, we cannot say with certainty which model is more valid. Moreover, the robustness of the random effects version over the fixed effects version gives us more reason to question Finn and Servoss's model. I conclude that out of all models tested in this study, the random effects Rasch model with binomial logistic regression of Model 2 is the most reliable and representative for evaluating school security. This conclusion is based on my own definition of school security, goodness-of-fit, and statistical normality and variance.

Perhaps adding more questions to the set in the analysis would yield more accurate results. Both this study and Finn and Servoss's study assume data is missing at random. However, omitting missing values may create inaccuracies in the analysis.

In the interpretation of these results, we should not confuse security with safety. It is possible that higher security creates safer schools, but it may also be true that schools may implement more security measures because they are located in unsafe neighborhoods or have had dangerous incidences occur in the past.

There may also be variation in the ways school administrators understand and answer questions in the survey. For example, in q38c, the survey asks if the school closes the campus for "most students during lunch." Answering this type of question may involve some degree of subjective speculation of how many is "most" and estimating the number of students who are allowed to leave. Other questions in this survey may also be subject to survey respondent bias. School administrators may tend to answer questions in a way that presents their school as highly secure.

Moreover, this study is subject to the bias of researchers. We have just shown that the selection of the questions changes the results of the analysis. Researchers act on bias when they choose questions to analyze based on what they believe demonstrates security in schools. Additionally, researchers' definitions of security may differ when there are multiple types of safety — physical, emotional, mental, etc.

Further research may be continued to improve these models and our understanding of school security. Models may be refined to handle data imputation rather than omitting missing values if data is missing not at random. Other modeling techniques may be tested, such as Bayesian approaches. Additionally, applications of this study and other aspects of security should be analyzed, such as students' feelings of safety within schools, frequency of occurrence of dangerous events in schools, and the relationship between school security and student learning.

Researchers should be cautious of associating school security with certain demographics. Consequences of inaccurate analyses may create or reinforce racial and socioeconomic assumptions and stereotypes. On the other hand, accurate results may inform people on focus areas for social, security, and educational improvement.

References

Perumean-Chaney, Suzanne E.; & Sutton, Lindsay M. (2013). Students and Perceived School Safety: The Impact of School Security Measures. *American Journal of Criminal Justice, 38(4),* 570-588.

Finn, Jeremy D. and Servoss, Timothy J. (2014) "Misbehavior, Suspensions, and Security Measures in High School: Racial/Ethnic and Gender Differences," *Journal of Applied Research on Children: Informing Policy for Children at Risk*: Vol. 5: Iss. 2, Article 11. Available at: http://digitalcommons.library.tmc.edu/childrenatrisk/vol5/iss2/11

Stats 485 Unit 3 Appendix

Emily MacQuarrie

4/26/2019

Load Libraries

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(lme4)
```

```
## Loading required package: Matrix
```

```
# suppress warnings
options(warn=-1)
```

Load Data

```
set.seed(485)
security_wide <- read.csv('http://dept.stat.lsa.umich.edu/~bbh/s485/data/security_wide.c
sv')
nschools <-nlevels(security_wide$school)
# Create dataframe for analysis
fs_items <- c(paste0('q38',c('c', 'd', 'h', 'g','n')), 'q40a', 'q38f')
security_wide_fs <- security_wide[c('school', fs_items)]
```

Analyze Complete Cases

```
# number of no's and yes's from schools
security_wide_fs %>% complete.cases() %>% table()
```

```
## .
## FALSE  TRUE
##    19   637
```

```
security_wide %>% dplyr::select(q38c, q38d, q38h, q38g, q38n, q40a, q38f) %>% sapply(function(vec)table(vec, exclude=NULL)) %>% knitr::kable()
```

	q38c	q38d	q38h	q38g	q38n	q40a	q38f
0	637	581	549	495	340	235	371
1	18	69	102	153	311	418	284
NA	1	6	5	8	5	3	1

```
# set index for schools based on number of yes responses
sum_index <- security_wide_fs %>% na.omit() %>% dplyr::select(-school) %>% rowSums()
table(sum_index)
```

```
## sum_index
##   0   1   2   3   4   5   6
##  76 153 182 130  71  22   3
```

Format Dataframe

```
security_tall_fs0 <-security_wide_fs %>% na.omit() %>% tidyr::gather("item", "response",
-school)
head(security_tall_fs0, 3)
```

```
##   school item response
## 1 id1011 q38c        0
## 2 id1021 q38c        1
## 3 id1022 q38c        0
```

Rasch Model

```
# binomial logistic rasch model
rasch0 <-glm(response~item+school -1,family=binomial, data=security_tall_fs0)
school_coef_names <-paste0("school",levels(security_wide$school))
# display number of yes's and no's from schools
table(school_coef_names %in% names(coef(rasch0)) )
```

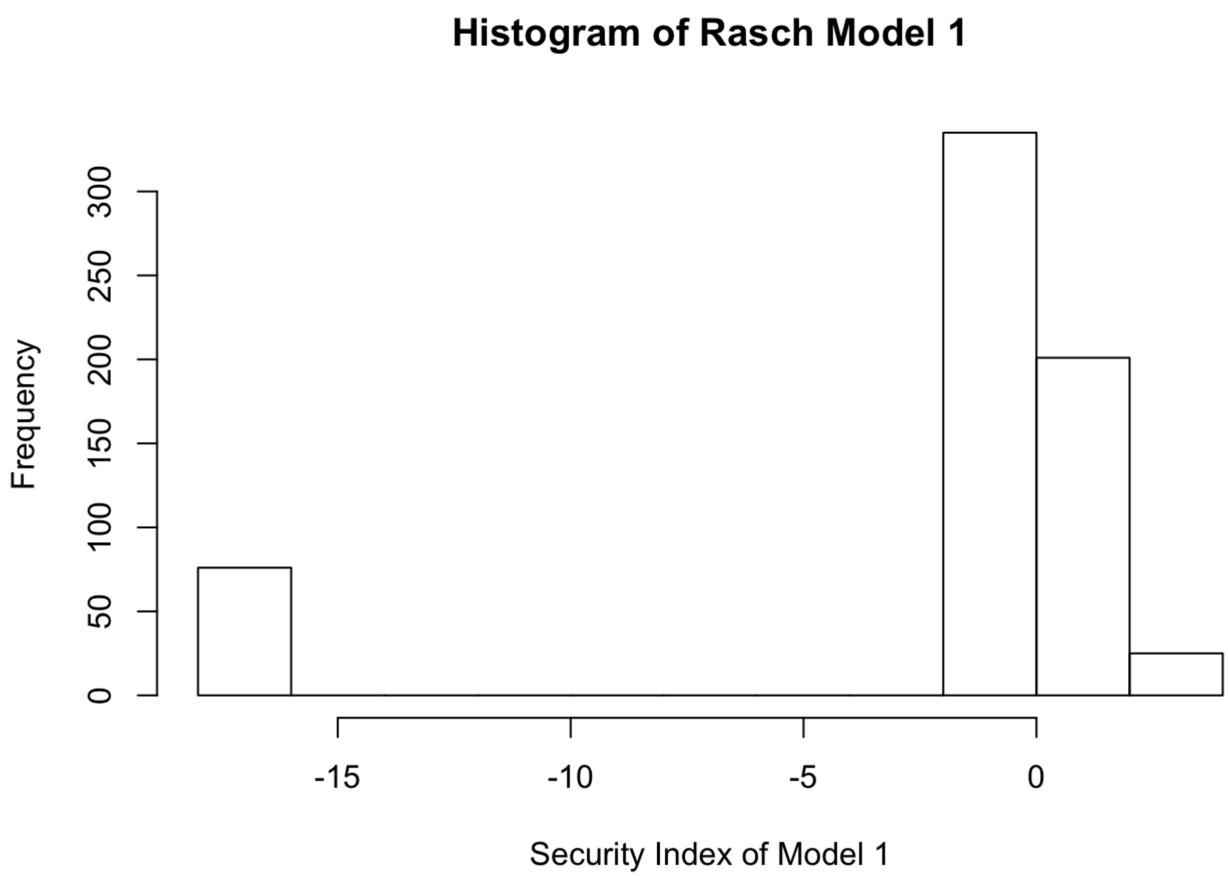
```
##
## FALSE  TRUE
##    20   636
```

Analyze Rasch Indices

```
rasch0_index <-c(schoolid1011=0,coef(rasch0)[-(1:7)])
table(round(rasch0_index))
```

```
##
## -18  -1   0   1   2   3   4
##  76 153 182 130  71  22   3
```

```
hist(round(rasch0_index), main="Histogram of Rasch Model 1", xlab="Security Index of Mod
el 1")
```




```
rank <- sort(rasch0_index, decreasing=TRUE)
head(rank)
```

```
## schoolid3741 schoolid3551 schoolid2402 schoolid2461 schoolid1951
##     4.109886     4.109886     4.109886     2.835690     2.835690
## schoolid1311
##     2.835690
```

Rasch Model with Different Survey Questions

```
# select questions for analysis
fs_items_c <- c(paste0('q38',c('a', 'e', 'g','h', 'm')), 'q40c', 'q40d')
security_wide_fs_c <- security_wide[c('school', fs_items_c)]
```

Analyze Complete Cases

```
security_wide_fs_c %>% complete.cases() %>% table()
```

```
## .
## FALSE  TRUE
##    21   635
```

```
sum_index_c <- security_wide_fs_c%>% na.omit()%>%dplyr::select(-school)%>% rowSums()
table(sum_index_c)
```

```
## sum_index_c
##   0   1   2   3   4   5   6   7
##  13  40 109 187 164  94  25   3
```

Format Dataframe

```
# reformat dataframe
security_tall_fs0_c <-security_wide_fs_c%>% na.omit()%>%tidyr::gather("item", "response"
,-school)
head(security_tall_fs0_c, 3)
```

```
##   school item response
## 1 id1011 q38a        1
## 2 id1021 q38a        1
## 3 id1022 q38a        1
```

Rasch Model

```
rasch0_c <-glm(response~item+school -1,family=binomial, data=security_tall_fs0_c)
head(coef(rasch0_c), 10)
```

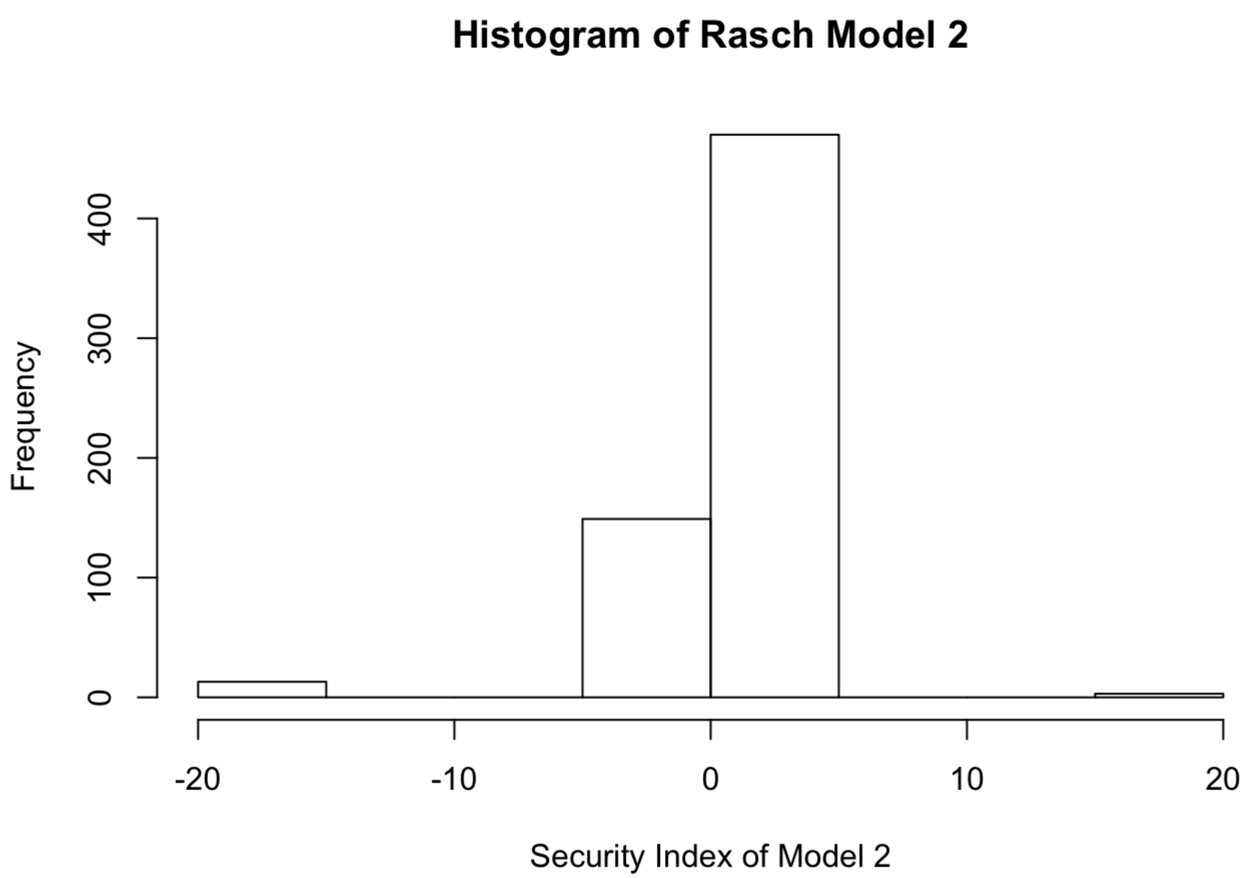
```
##     itemq38a     itemq38e     itemq38g     itemq38h     itemq38m
##   0.07736402  -0.25942517  -2.88349723  -3.52644941  -2.01700089
##     itemq40c     itemq40d schoolid1021 schoolid1022 schoolid1031
##   1.29210496  -2.72201863   2.92436550   1.04541744  19.56177436
```

Analyze Rasch Indices

```
# set school security indices from Rasch model with selected subset of questions
rasch0_index_c <-c(schoolid1011=0,coef(rasch0_c)[-(1:7)])
table(round(rasch0_index_c))
```

```
##
## -17  -1   0   1   2   3   4  20
##  13  40 109 187 164  94  25   3
```

```
# display histogram of security indices
hist(round(rasch0_index_c), main="Histogram of Rasch Model 2", xlab="Security Index of M
odel 2")
```



Comparison of Models

```
# 629 elements in intersection
common_ids = sort(intersect(names(rasch0_index), names(rasch0_index_c)))
r0_trunc = rasch0_index[common_ids]
r1_trunc = rasch0_index_c[common_ids]
ranks0 = rank(r0_trunc, ties.method = "random")
ranks1 = rank(r1_trunc, ties.method = "random")
```

Plot Comparison of Ranks

```
# plot school rank from model 1 and model 2
# each school is a data point
ggplot(data.frame(cbind(ranks0, ranks1)),aes(ranks0, ranks1)) +
  geom_point() + ggtitle("Comparison of Model 1 to Model 2") +
  xlab('Model 1 Ranks') + ylab('Model 2 Ranks')
```

Correlation of Model Ranks

```
# 1 indicates correlation
cor(ranks0, ranks1, method ="spearman")
```

```
## [1] 0.5507364
```

```
cor(ranks0, ranks1, method ="kendall")
```

```
## [1] 0.3850111
```

The two models are closer to no correlation. The FS model is unreliable but valid.

Compare Fixed Effects to Random Effects

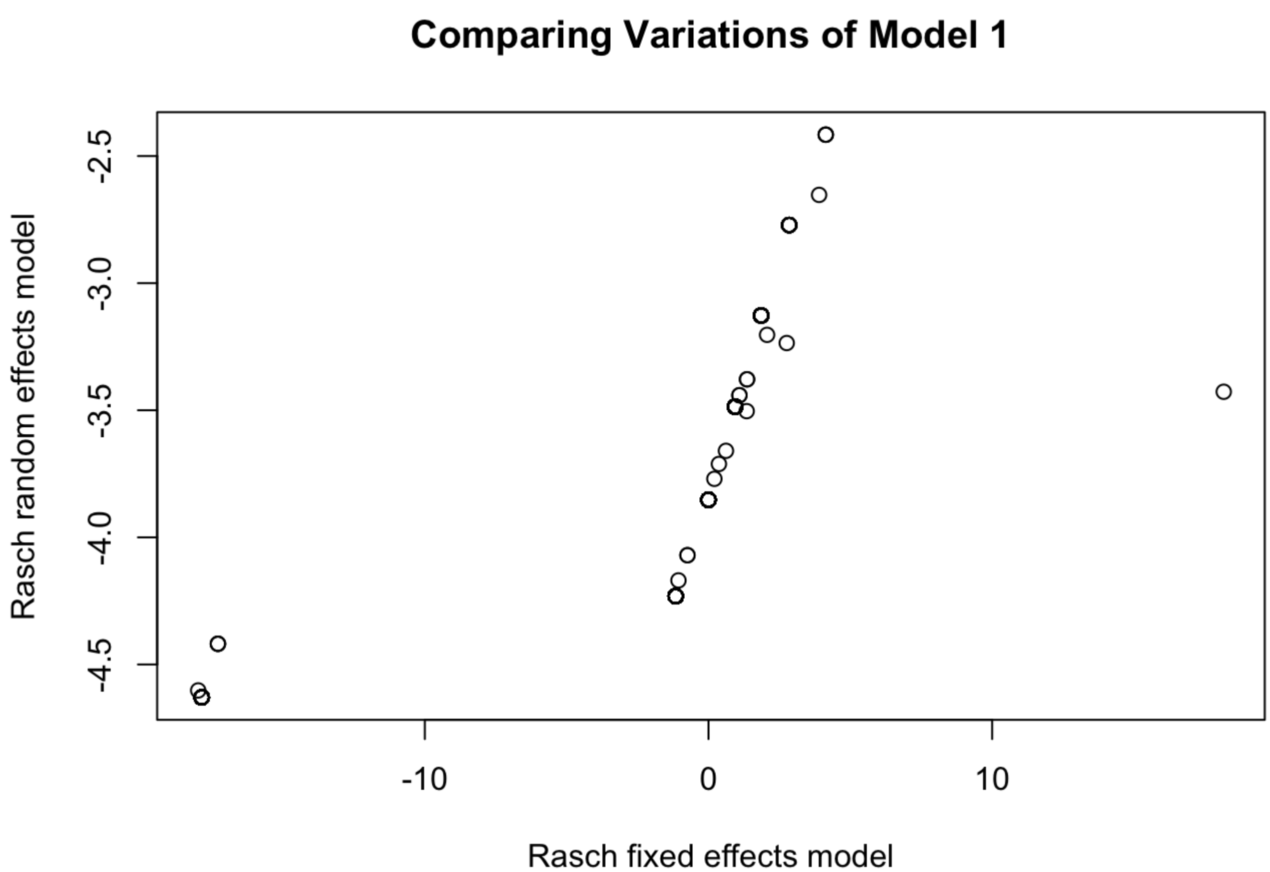
```
security_tall_fs1 <-security_wide_fs%>%tidyr::gather("item", "response",-school)
rasch01 <-glm(response~item+school -1,family=binomial, data=security_tall_fs1)
rasch01_index <- c(schoolid1011=0,coef(rasch01)[-(1:7)])

rasch1 <- glmer(response~item+(1|school),family=binomial, data=security_tall_fs1)
head(coef(rasch1)$school, 3)
```

```
##        (Intercept) itemq38d itemq38f itemq38g itemq38h itemq38n itemq40a
## id1011   -3.852357 1.486856 3.544458 2.520415 1.963539 3.749774  4.50963
## id1021   -2.771385 1.486856 3.544458 2.520415 1.963539 3.749774  4.50963
## id1022   -4.629396 1.486856 3.544458 2.520415 1.963539 3.749774  4.50963
```

```
# security indices are now intercept values
rasch1_index <-coef(rasch1)$school[,"(Intercept)"]
sum(security_wide$school != row.names(coef(rasch1)$school))
```

```
## [1] 0
```

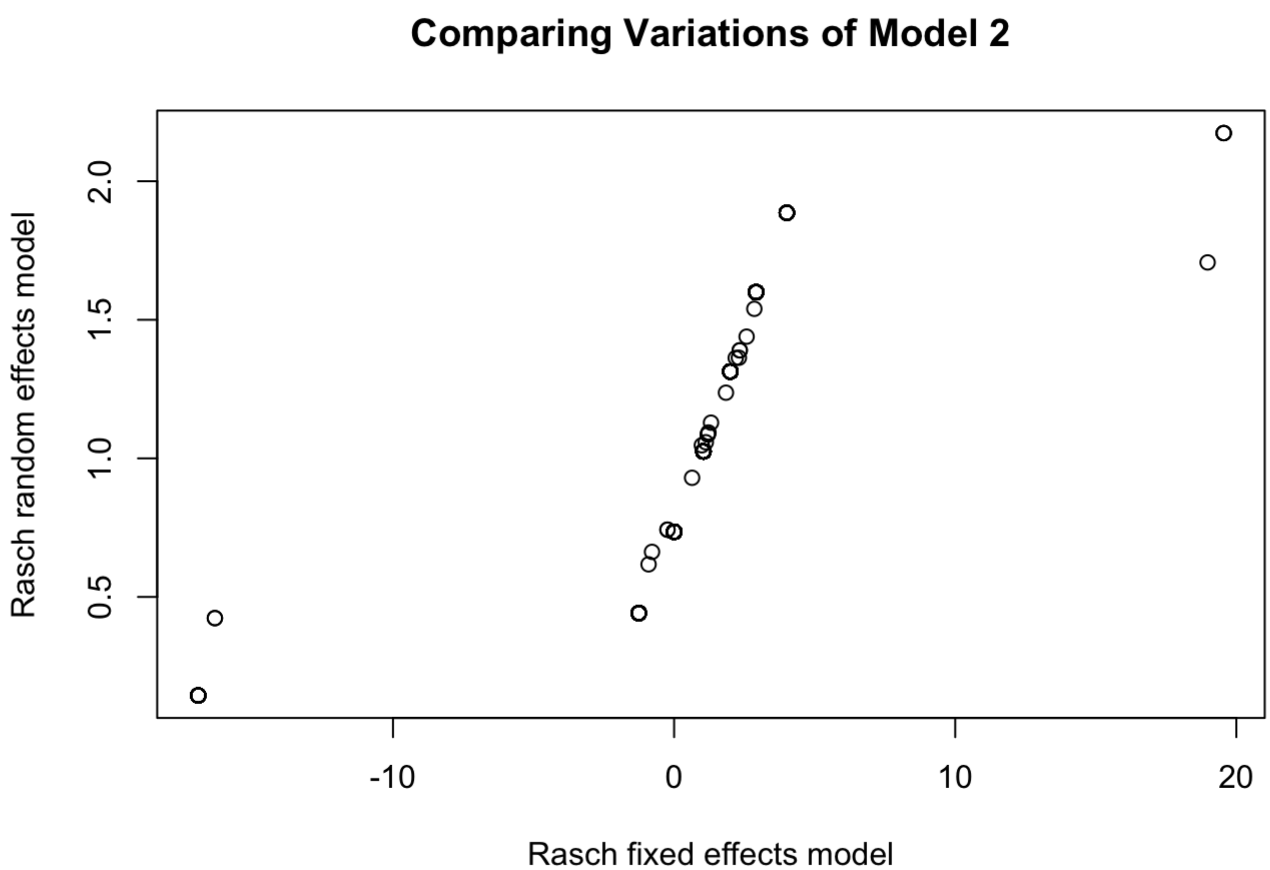
```
plot(rasch01_index, rasch1_index, main = "Comparing Variations of Model 1", xlab="Rasch
 fixed effects model",ylab="Rasch random effects model")
```



```
security_tall_fs1_c <-security_wide_fs_c%>%tidyr::gather("item", "response",-school)
rasch01_c <-glm(response~item+school -1,family=binomial, data=security_tall_fs1_c)
rasch01_index_c <- c(schoolid1011=0,coef(rasch01_c)[-(1:7)])

rasch1_c <-glmer(response~item+(1|school),family=binomial, data=security_tall_fs1_c)
head(coef(rasch1_c)$school, 3)
```

```
##        (Intercept)   itemq38e  itemq38g  itemq38h  itemq38m  itemq40c
## id1011   0.7348392 -0.2177757 -2.400306 -2.943407 -1.683904 0.9771565
## id1021   1.5998726 -0.2177757 -2.400306 -2.943407 -1.683904 0.9771565
## id1022   1.0251529 -0.2177757 -2.400306 -2.943407 -1.683904 0.9771565
##         itemq40d
## id1011 -2.267218
## id1021 -2.267218
## id1022 -2.267218
```

```
rasch1_index_c <-coef(rasch1_c)$school[,"(Intercept)"]
plot(rasch01_index_c, rasch1_index_c, main = "Comparing Variations of Model 2",xlab="Ras
ch fixed effects model",ylab="Rasch random effects model")
```

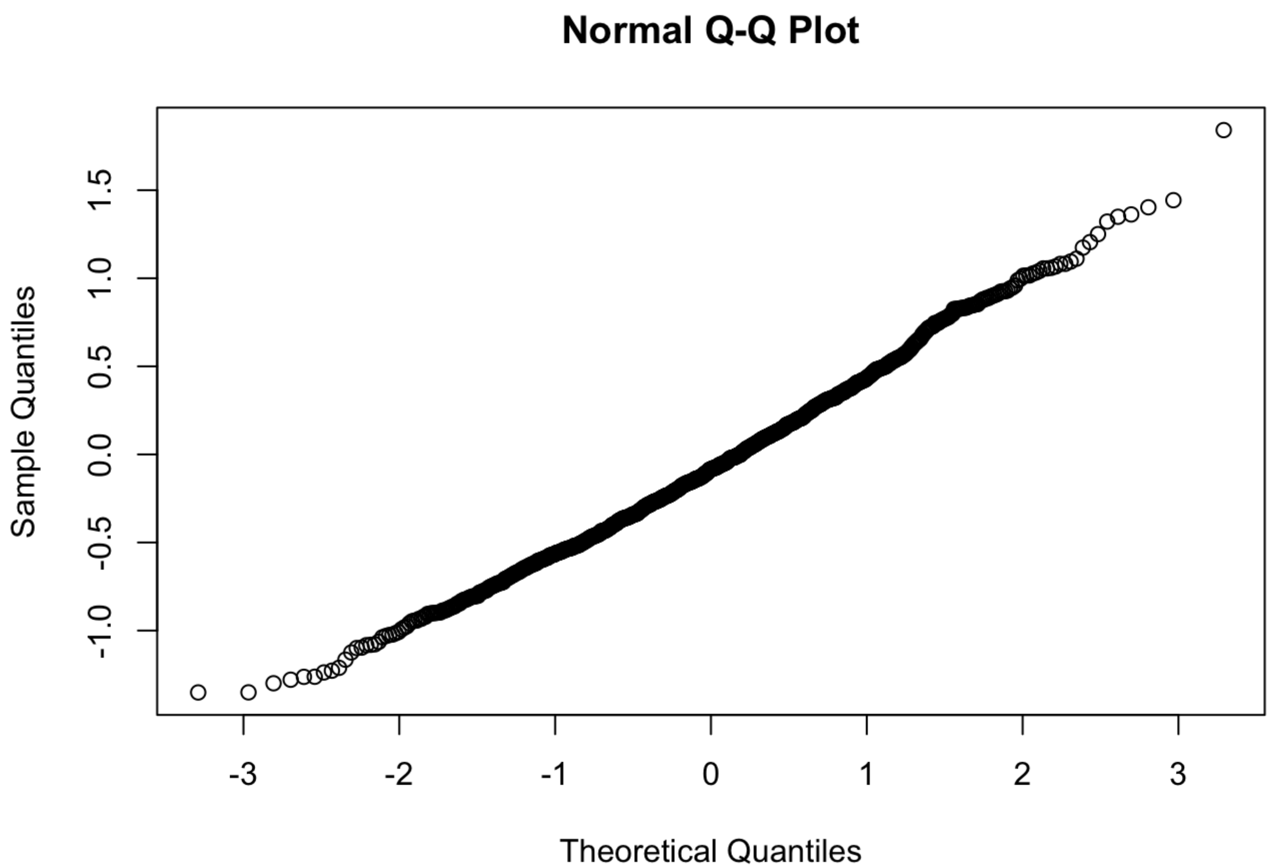


Bootstrap for random effects personalized model

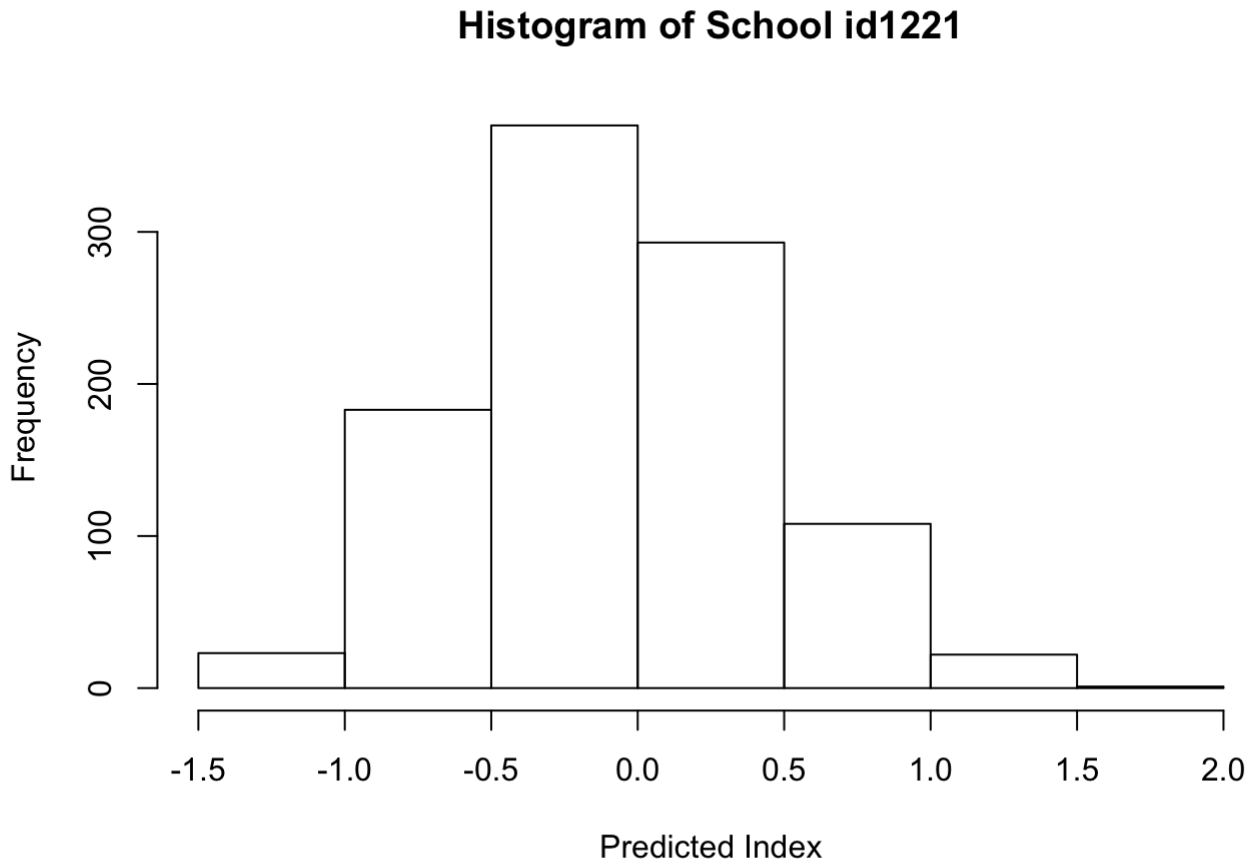
```
bootreps = 1000
indices = sample(1:nrow(security_tall_fs1_c), size = 3)
ids = as.character(security_tall_fs1_c[, "school"][indices])
# Printing school ids to analyze
ids
```

```
## [1] "id1221" "id2683" "id3941"
```

```
rboot = function(statistic, simulator, B)
{
  tboots = replicate(B, statistic(simulator()))
  if(is.null(dim(tboots)))
  {
    tboots = array(tboots, dim = c(1, B))
  }
  return(tboots)
}

resampler = function()
{
  whichrows = sample(1L:nrow(security_tall_fs1_c), nrow(security_tall_fs1_c), replace =
T)
  security_tall_fs1_c[whichrows, ]
}

coefs = function(bootspl)
{
  refitted.mod = update(rasch1_c, data = bootspl)
  schools = coef(refitted.mod)$school
  vals = schools[, "(Intercept)"]
  names(vals) = row.names(schools)
  unlist(vals)
}
boot.stats <- rboot(coefs, resampler, bootreps)

vec1 = numeric(length = bootreps)
vec2 = numeric(length = bootreps)
vec3 = numeric(length = bootreps)
for (i in 1:bootreps)
{
  vec1[i] = unlist(boot.stats[i])[ids[1]]
  vec2[i] = unlist(boot.stats[i])[ids[2]]
  vec3[i] = unlist(boot.stats[i])[ids[3]]
}
vec1 = vec1[!is.na(vec1)]
vec2 = vec2[!is.na(vec2)]
vec3 = vec3[!is.na(vec3)]

#Test for normality and variances
qqnorm(vec1)
```




```
hist(vec1, main = paste("Histogram of School" , ids[1]), xlab = "Predicted Index")
```



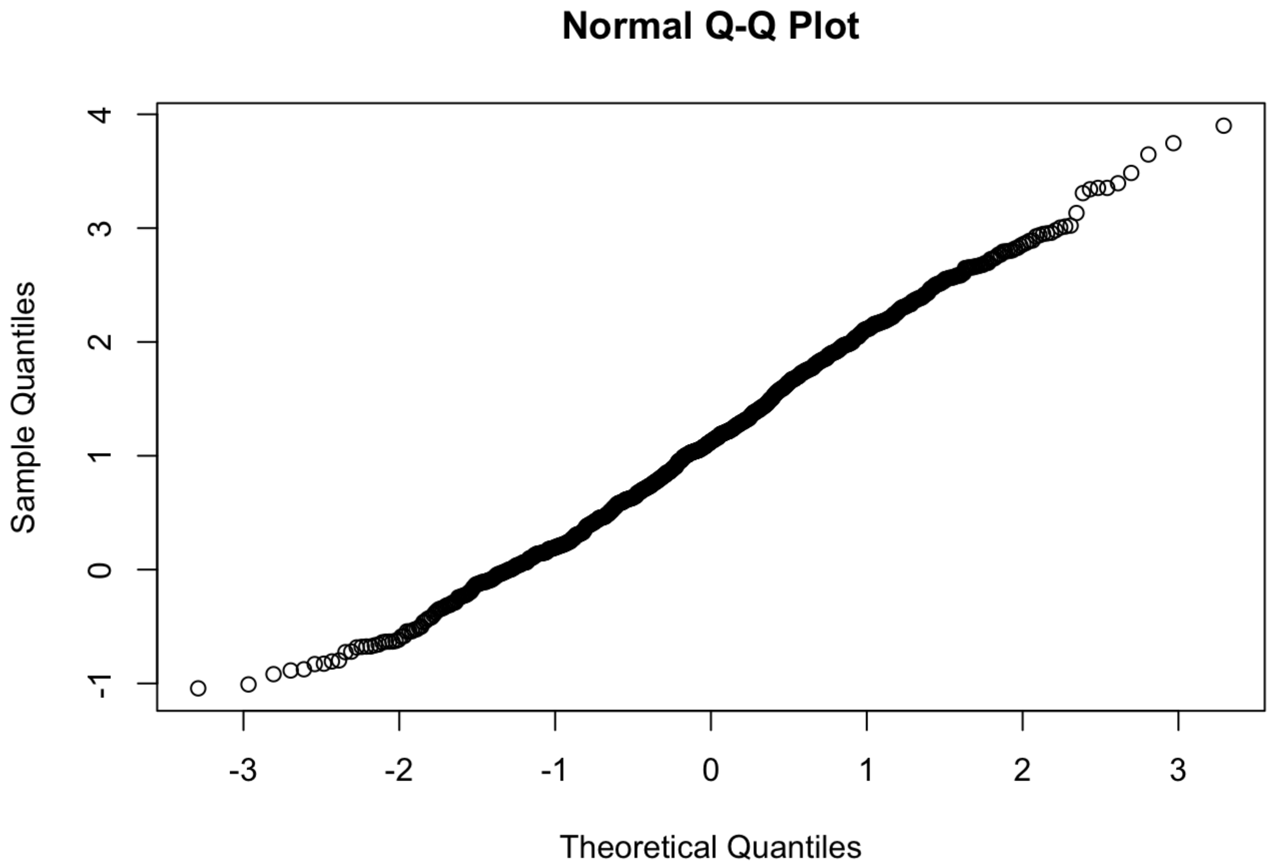

```
# Variance of bootstrapped indices
var(vec1)
```

```
## [1] 0.255186
```

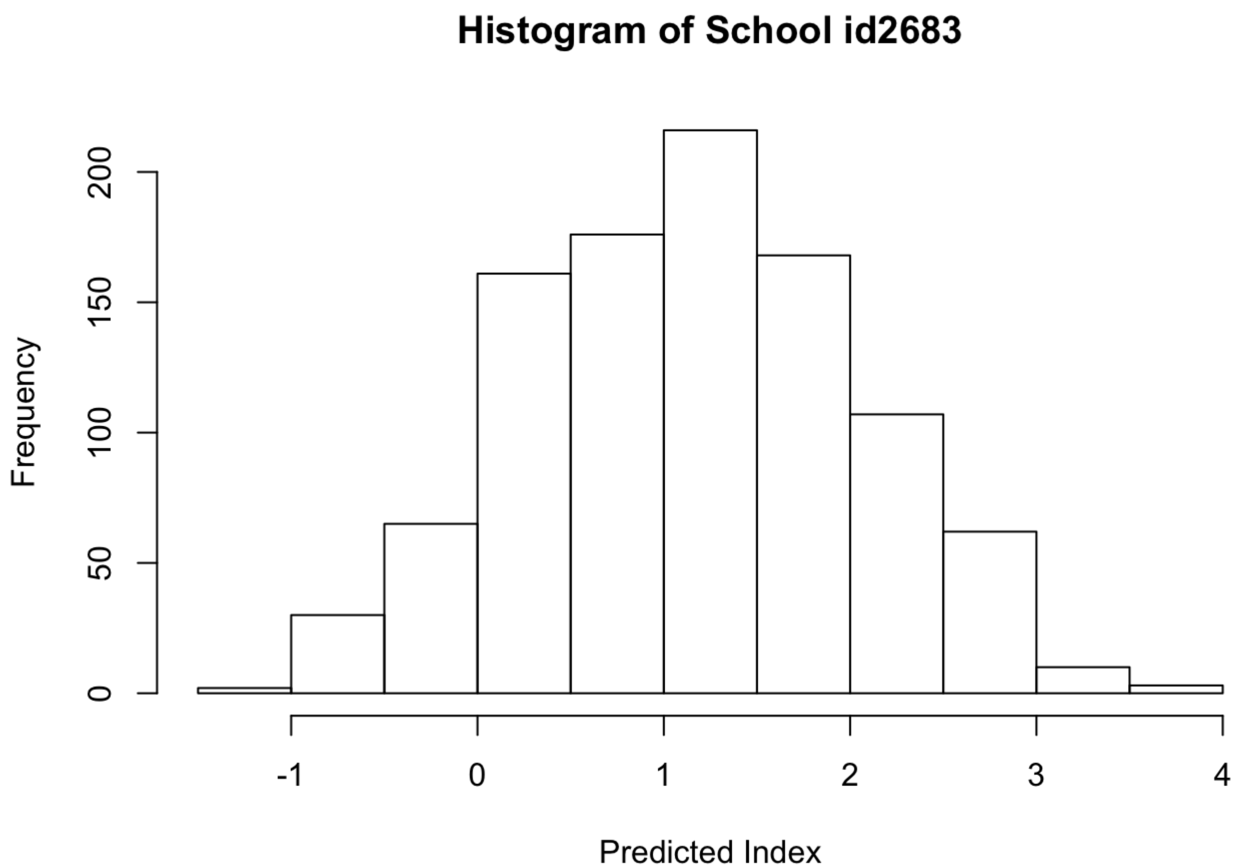
```
# Calculate sample sd of residuals
boot_truth <- coef(rasch1_c)$school[ids[1], "(Intercept)"]
sd_1 <- sd(vec1 - boot_truth)
sd_1
```

```
## [1] 0.5051594
```

```
qqnorm(vec2)
```




```
hist(vec2, main = paste("Histogram of School" , ids[2]), xlab = "Predicted Index")
```



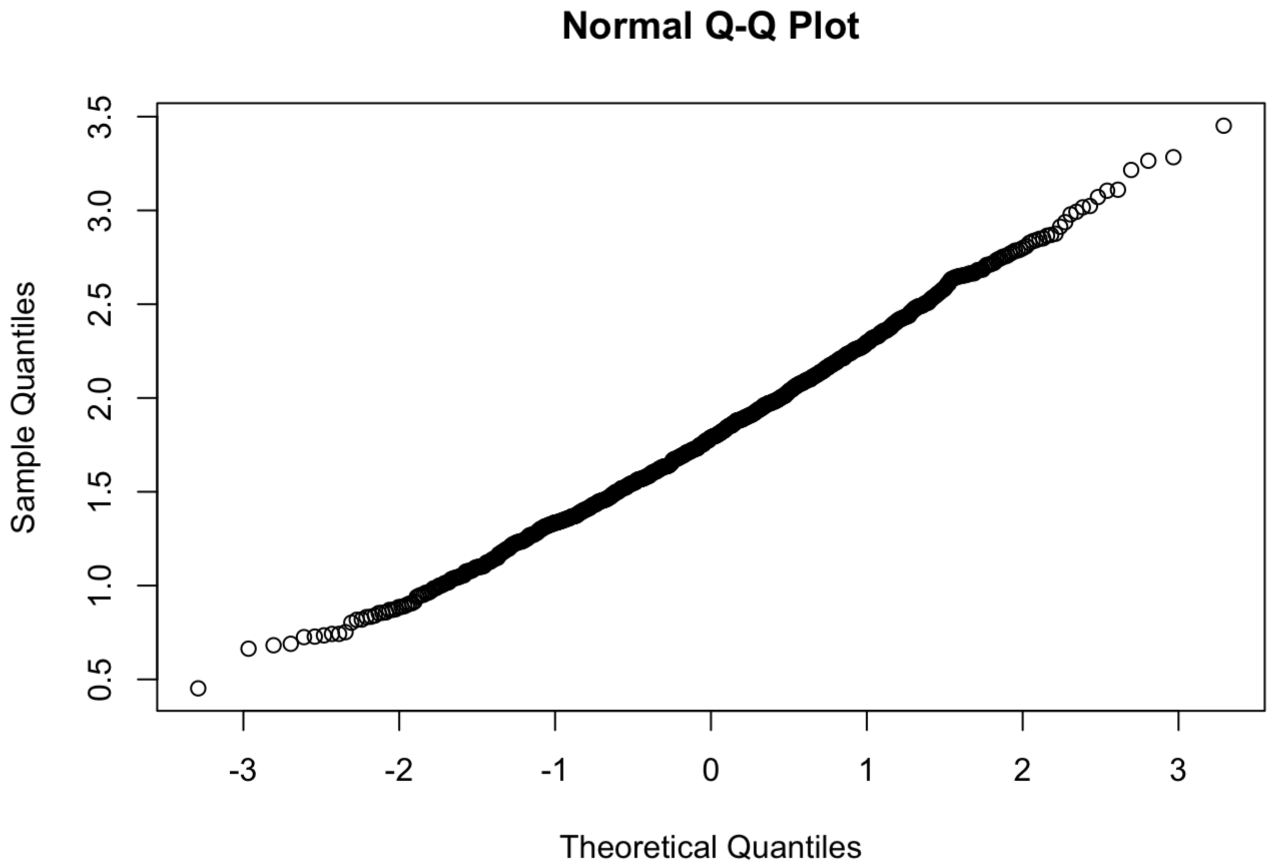

```
# Variance of bootstrapped indices
var(vec2)
```

```
## [1] 0.8017769
```

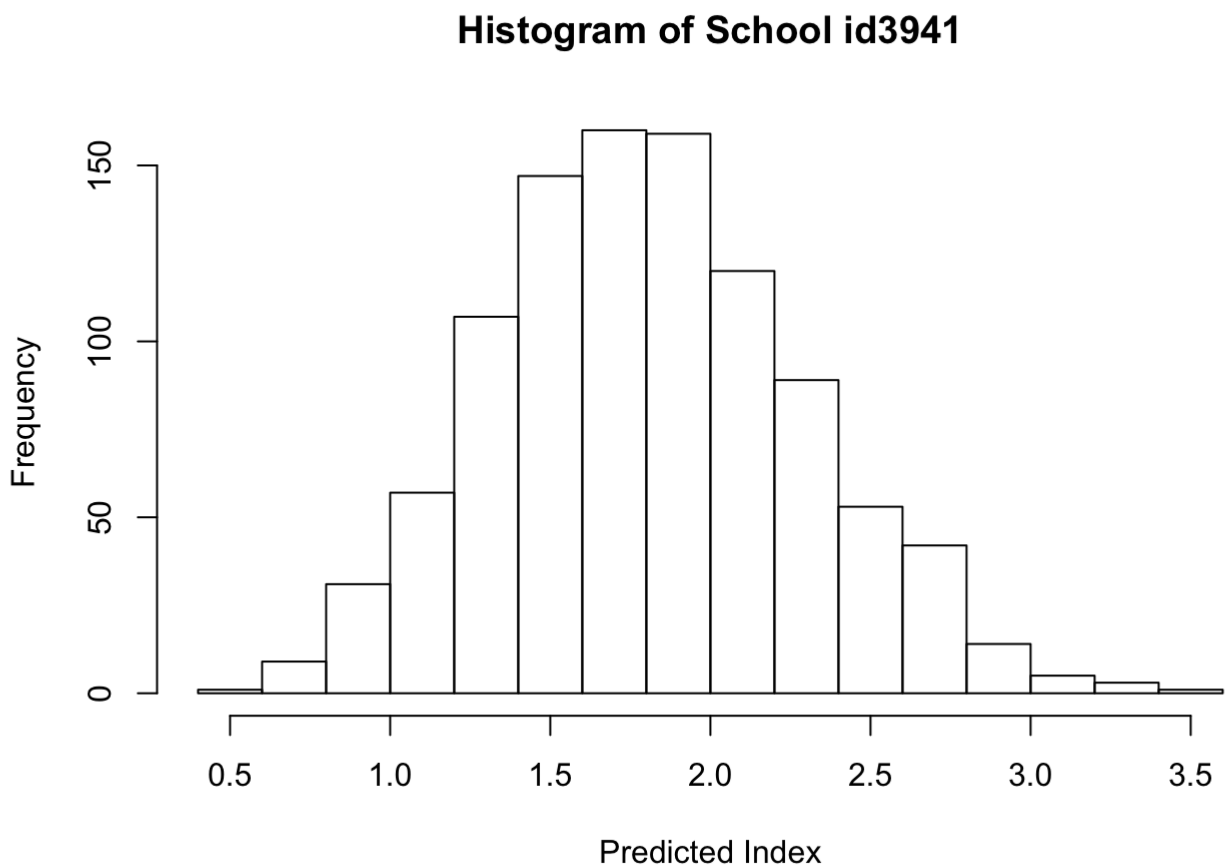
```
# Calculate sample sd of residuals
boot_truth <- coef(rasch1_c)$school[ids[2], "(Intercept)"]
sd_2 <- sd(vec2 - boot_truth)
sd_2
```

```
## [1] 0.89542
```

```
qqnorm(vec3)
```




```
hist(vec3, main = paste("Histogram of School" , ids[3]), xlab = "Predicted Index")
```



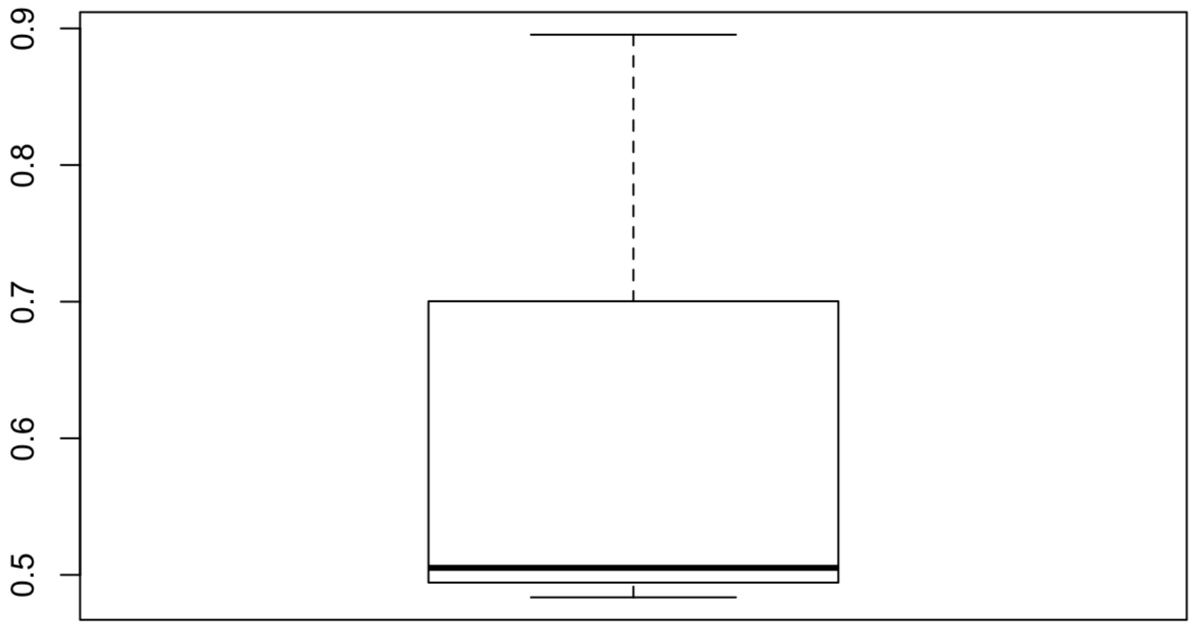

```
# Variance of bootstrapped indices
var(vec3)
```

```
## [1] 0.2338495
```

```
# Calculate sample sd of residuals
boot_truth <- coef(rasch1_c)$school[ids[3], "(Intercept)"]
sd_3 <- sd(vec3 - boot_truth)
sd_3
```

```
## [1] 0.4835798
```

```
boxplot(c(sd_1, sd_2, sd_3))
```




```
avgsd <- mean(c(sd_1, sd_2, sd_3))
avgsd
```

```
## [1] 0.6280531
```

```
# difference between average residual sd and residual sd for each school
c(sd_1, sd_2, sd_3) - avgsd
```

```
## [1] -0.1228937  0.2673669 -0.1444732
```