Categorical analysis on SOTES data

Presented to

Dr. Lee

Department of Mathematics and Statistics

San José State University

In Partial Fulfillment

Of the Requirements for the Class

MATH 258

By

Jung-a Kim

December, 2019

Student Opinion of Teaching Effectiveness has been widely used for measuring the performance of instructors for the promotion in universities. However, its validity for assessing the instructor's teaching has been controversial over the last few years and it is deemed responsible for grade inflation which discourages critical thinking development and academic achievement for students. SOTES process summarizes to submitting one ordinal scale score as the rating of the instructor which may lack various aspects of teaching skills and it is in doubt that the score reflects the purpose of the survey.

The data used in this analysis was collected over a year during Fall 2016 and Spring 2017 which includes variables that potentially influence the student's tendancy to rate instructors. The response variable, the rating of the instructor is ordinally scaled from 1 to 5("Very ineffective", "Ineffective", "Somewhat Effective", "Effective", "Very Effective"). There are 86,170 observations with the variables including college, department, class size, course level, student level, expected grade, and official grade.

It is easy to observe that students tend to to submit high rating scores for the instructors who are lenient towards grading. However, the survey is collected one month before the grade release and yet half of the ratings are 5. This makes us wonder whether the students already know that they will receive high grades or there are some other reasons to explain the phenomenon.

The analysis will first explore some of the characteristics of the students related to the ratings of the instructors and discuss a generalized linear model to discover the relationships between them.

The missing data in each variable were excluded in this analysis and it is solely based on undergraduate and graduate students who take courses with letter grade system. The numerical values used in each tables and figures are rounded by 2 or less. The confidence intervals at each occurence are asymptotic normal confidence intervals at 0.05 significance level.

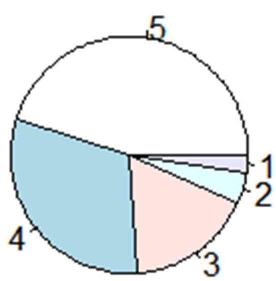


Figure 1. Proportion of rating

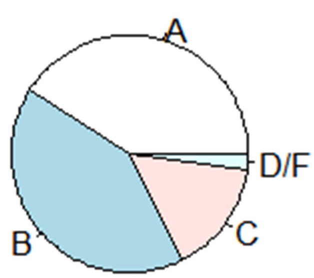


Figure 2. Proportion of expected grade

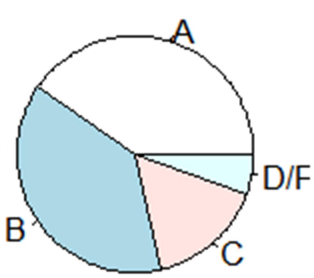


Figure 3. Proportion of actual grade

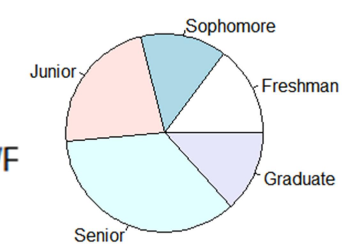


Figure 4. Proportion of student levels

From figure 1, 45% is 'Very Effective(5)' and 76% of the rating is at least 'Effective(4)' which means the rating is inflated. An instructor would not be very surprised if s/he received the rating score 5. Considering that instructors compete for promotion, it may be imperative for them to get more rating score 5 than others.

From figure 2, expected grade has analogous distribution to rating as it had been suspected. 83% of the students expect to receive at least 'B' and 42% of the students expect to receive 'A' as their final grades. Actual grades in figure 3 looks almost the same as the expected grades in figure 2.

From figure 4, there are noticeably more seniors than other student levels. About half of the students are either senior or graduate students who may care about the grades even more since they are closer to the stage of preparing for employment than the lower level students.

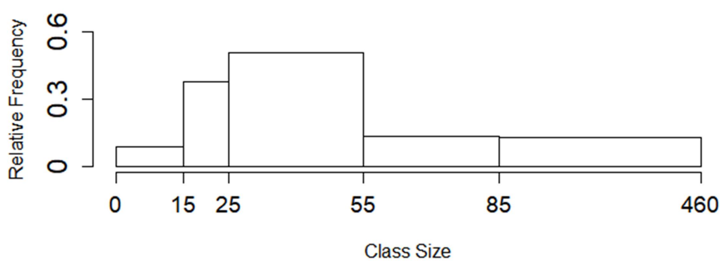


Figure 5. Proportions of categorized class sizes

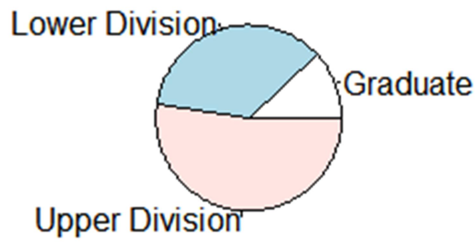


Figure 6. Proportions of course levels

In general, the rating is negatively correlated with class size. The class sizes are categorized into 5 groups where the distributions are similar within each range.

Most of the classes are sized between 15 and 55. Extremely large classes are usually courses required for all majors so there are fewer number of them than major courses. Small classes are usually in the setting of project or thesis which involves supervision of the instructor rather than teaching skills.

The proportions of course levels correspond to the proportions of student levels. For example, 13% of the courses are graduate level and the same proportion of SJSU students are graduate students. This means that the courses are designed to corporate students of each level considering their proportions during the year.

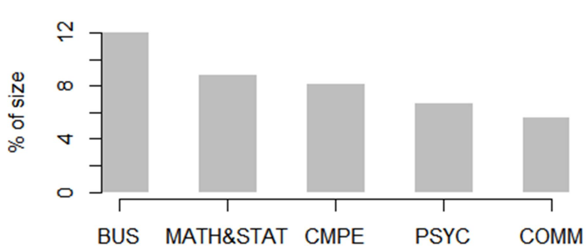


Figure 7. Proportions of 5 largest departments

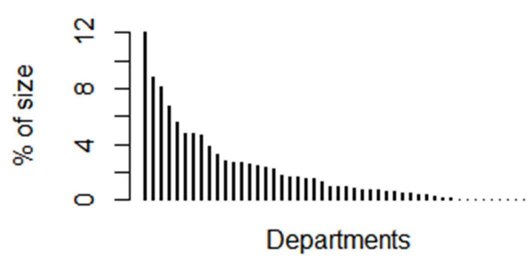


Figure 8. Proportions of all departments

There are 47[1] departments in SJSU and particular departments such as business, mathematics & statistics, computer engineering, psychology, and communication studies have higher proportion

[1] Business includes Business 1 through Business 5. Credential departments Math Education(MTED) and Social Science Education(SSED) are excluded because the credential students are excluded from this analysis.

than the other departments. In general, the number of assistant professors is proportional to the size of the departments. Thus, professors in large departments would have more competitors for tenure promotion than those in relatively small departments.

60% of the students were enrolled in Fall which is to be expected since new students are enrolled in Fall and students usually graduate in Spring.

Expected grade	Rating of the instructors				
	Very effective(5)	Effective(4)	Somewhat effective(3)	Ineffective(2)	Very ineffective(1)
A	80.4	-27.0	-50.2	-27.5	-20.2
B	-29.1	30.3	5.2	-3.6	-5.1
C	-58.1	-0.4	53.0	29.1	21.2
D/F	-28.7	-10.3	19.7	33.0	32.8
$\chi^2 = 12401.64 > 21.03 = \chi^2_{0.05;12}$					

Table 1. Standardized Pearson residuals of the contingency table between expected grades and ratings of the instructors

The first suspicion about the association between students' expected grades and their rating of the instructors can be verified by Pearson's Chi-square test of independence. The test compares the frequency in each cell of the contingency table to the frequency under independence assumption.

The highlighted residuals serve as evidence that there is strong association between expected grades and rating of the instructors. Compared to independence assumption, there are more A-expectators[2] who submit score 5, there are more B-expectators that submit score 4 and so on. This indicates there is monotonically positive association between expected grades and rating of the instructors. This association is particularly more extreme for A-expectators. One following question may be whether this association depends on the accuracy in students' prediction.

Exp-ected grade	Actual grade			
	A	B	C	D/F
A	30	10	1	0
B	10	23	7	1
C	1	5	8	3
D/F	0	0	1	1

Table 2. Proportion(%) of expected grade vs. actual grade

Expected grade	A-student's Rating	
	5	5 or lower
A	46	28
B	10	16

Table 3. Proportion(%) of expected grade vs. A-student's rating

Estimation of grade	A-expectator's Rating	
	5	5 or lower
Over-estimation	18	10
Correct estimation	45	28

Table 4. Proportion(%) of expected grade vs. A-expectator's rating

62% of the students correctly predicted their letter grades, 22% overestimated their grades, and 17% underestimated their grades as in table 2. The positive association between expected grade and rating did not change whether they overestimated, accurately estimated, or underestimated. As one summary statistics, Goodman-Kruskal gammas are 0.48, 0.45, and 0.52 respectively.

[2] Students who anticipate receiving 'A' as the final grade

In table 3, the probability that A-students[3] who made correct prediction submits score 5 is 62%(= 46/(46+28)). On the other hand, the probability that A-students who expected 'B' submits score 5 is about 40%(= 10/(10+16)). This indicates that if an A-student is wrong in their prediction, it will lower the chance to rate 5 by 22~23% based on 95% wald confidence interval.

From table 4, students who expected 'A' that they did not deserve are about 1.5~2.7% more likely to submit score 5 than students who expected 'A' that they deserved. This indicates that students who overestimate themselves are more likely to submit score 5 than realistic A-students.

Overall, students who overestimate themselves to receive 'A' are more likely to submit score 5 than actual A-students who know that they deserve at least 'B'. The following analysis shows which type of students overestimate themselves.

Student level	Estimation of grade: Over-estimation	Accurate	Under-estimation
Freshman	4.3	-0.7	-3.9
Sophomore	3.1	-2.4	-0.3
Junior	2.4	-3.2	1.6
Senior	-5.2	-1.8	8.0
Graduate	-3.3	9.8	-9.1
$G^2 = 190.52 > 15.51 = \chi^2_{0.05;8}$			

Table 5. Standardized Pearson residuals between student level and estimation of grades

Student level	Estimation of grade: Over-estimation	Accurate
Fresh.+ Soph.+ Jr.	16	43
Sr.	10	30

Table 6. Subtable(1) of table 5 with proportions(%)

Student level	Estimation of grade: Accurate	Over-estimation
Grad.	10	3
Under-grad.	64	23

Table 7. Subtable(2) of table 5 with proportions(%)

There is an evidence of association between student level and their estimation of final grades supported by the likelihood-ratio test of independence. The highlighted residuals in table 5 indicate that freshman, sophomore, and juniors are more likely to overestimate their grades than other students. It is noteworthy that seniors are more likely to underestimate their grades compared to other students and graduate students are best at predicting their grades. From one subtable(table 6), seniors are 1.6~2.3% less likely to overestimate their grades than freshmen, sophomores, and juniors. In another subtable(table 7), graduate students are 2.2~3.2% more likely to predict their grades accurately than undergraduates. The following partial table shows which type of students overestimate themselves to receive 'A'.

Student level	Overestimating group's expected grade: A	B	C
Graduate	-7.9	3.4	7.0
Senior	-11.5	6.1	8.7
Junior	-8.9	7.7	2.4
Sophomore	-3.5	5.0	-1.9
Freshman	37.8	-27.5	-17.5
$G^2 = 1776.54 > 15.51 = \chi^2_{0.05;8}$			

Table 8. Standardized Pearson residuals between student level and their expected grades within overestimating group

Student level	Overestimating group's expected grade: A	B
Freshman	13	1
The others	47	39

Table 9. Subtable of table 8 with proportions(%) within overestimating group (17821 students)

[3] Students who actually received 'A' as the final grade

Table 8 shows strong association between expected grades and student level within the group of students who overestimated their grades. The highlighted residual indicates that if a student overestimates himself to receive 'A', s/he is more likely to be a freshman than any other students. In fact, overestimating freshmen are 1.6 to 1.7 times more likely to expect 'A' than other overestimating students inferred from table 9.

To summarize what has been discovered so far, there is positive association between expected grade and rating. And students who overestimate themselves to receive 'A' are the most likely ones to rate 5 and they are most likely to be freshmen.

From figure 1, it was noted that receiving score any less than 5 is deterimental to an instructor's average rating. A generalized linear model can be used for predicting the probability of receiving score 5 using several relevant variables. The model for this data is a Bernoulli random variable with the probability of receiving score 5. A canonical link for this model is a logit link that generalizes the linear model assuming that log-odds of receiving score 5 has linear relationship with the variables.

The model[4] has 6 main effects and three two-way interaction effects. The main effects are 'semester', 'course level', 'department', 'student level', 'class size', and 'expected grade'. Each main effect has association with the rating which can be shown by the test of independence[5]. The variable 'college' had multicollinearity with 'department' since colleges are the partitions of departments. The actual final grade was not chosen as a predictor since the rating is submitted before the final grade is released such that it cannot possibly be used as a predictor that affects the rating. 'Class size' is specifically the number of enrolled students in a class. The original dataset had three variables related to 'class size', but the number of registered students were mostly the same with very high correlation of .99 with the number of enrolled students. The capacity of class also had high correlation of .96 with the number of enrolled students as expected since it is designed to accomodate the students that need to take the course in the semester. Student level and course level as main effects are positively correlated with rating[6].

The two-way interaction effects were selected in the model when adding one variable changed the other variable's coefficient by at least 0.1 which is the same as changing the odds ratio by at least 10%.

[4] The model specification is in 'SOTES_model_Jung-a Kim.pdf' file
[5] Refer to the appendix
[6] Test of independence and odds ratios are in the appendix

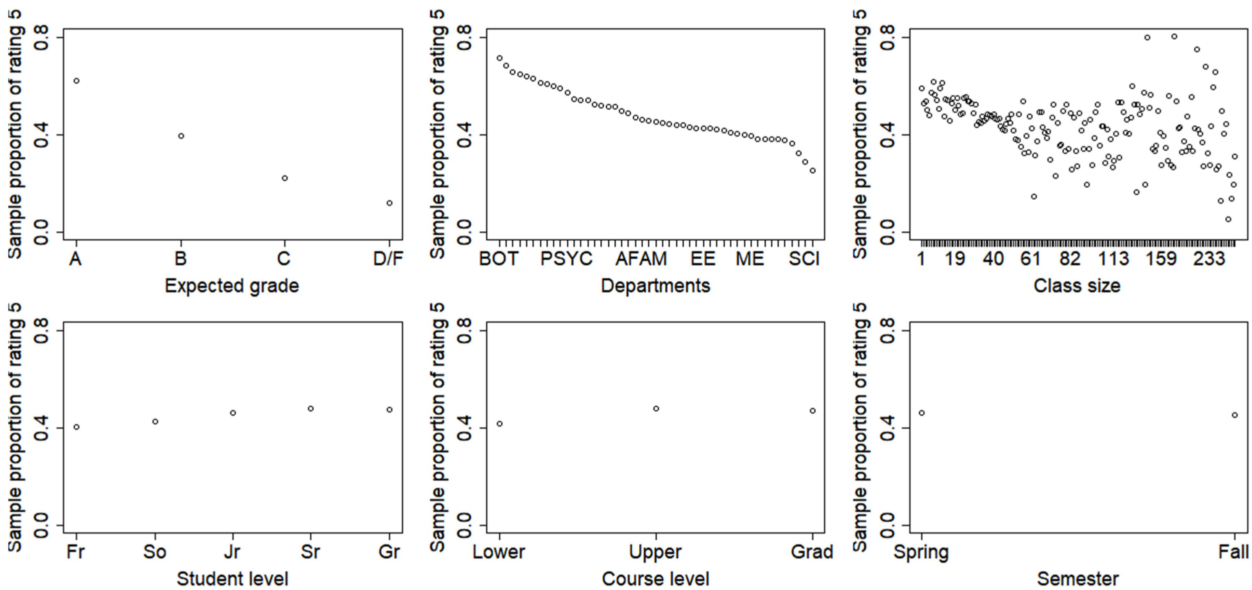


Figure 9. Sample proportion of Rating 5 against main effects

At one glance from figure 9, expected grade, departments, and class size have much stronger association with rating 5 compared to student level, course level, and semester. The expected grade has the strongest impact on rating. Rating also depends a lot on departments. Those who have high tendancy to rate 5 and those who have low tendancy to rate 5 are mostly from relatively small departments. As in figure 10, the sample probabilities from departments of size less than 1%[7] tend to vary a lot compared to large departments. Since small departments have fewer number of professors, this trend could directly indicate that ratings of professors depend highly on the department s/he belongs to.

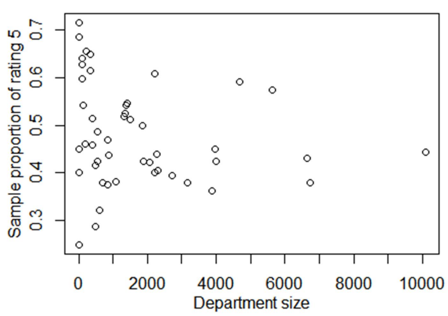


Figure 10. Sample proportion of Rating 5 against department size

Back to figure 9, there is a decreasing trend in rating as class size increases, but the ratings in classes with more than about 120 students tend to vary a lot. Among these large classes, many of them were departments from college of science and engineering which usually have low ratings[8].

Students in the fall or spring semester do not seem to show difference in rating 5. Student levels and course levels seem to have very weak positive association with rating 5, but the strength of the association depends a lot on departments as observed in the following interaction plots.

---

[7] Among the small departments, Botany, Zoology, SOCS, Astronomy, Humanities, Social Sciences(SOCS), Urban & Regional Planning, Sociology(SOCI), and Global Studies students tend to rate higher while Entomology, Science Education Program, Technology, Materials Engineering, American Asian Studies, Aviation, and Biomedical Engineering students tend to rate lower.

[8] The proportions are in the appendix

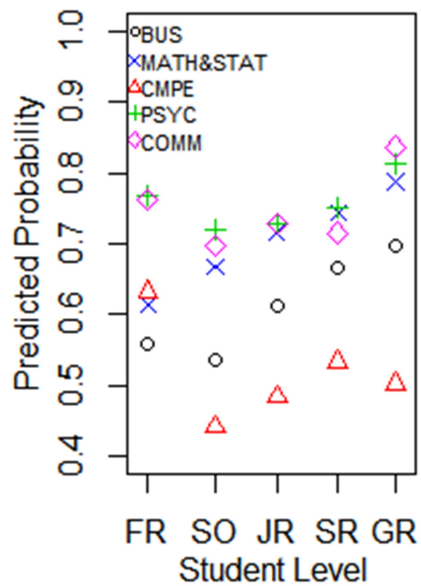
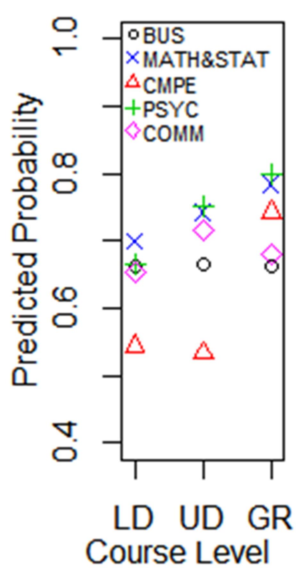


Figure 11. Predicted probability of rating 5 when a student takes an upper division class of 25 people in Fall and expects 'A'

One of the interaction effects is between student level and department. For a meaningful comparison, 25 sample students with common levels of the remaining variables were created who take an upper division course with class size of 25 in the fall semester and expect grade as 'A'. The students are in different academic levels belonging to different departments. Figure 11 shows five most common departments which account for 41% and how they affect the rating differently according to each student level. For example, there is monotonic positive association between mathematics and statistics' student level and rating. Business students and computer engineering students show a similar pattern of rating where sophomores are the most reluctant to rate 5 while psychology students and communication students show a similar u-shape pattern as students mature. It is noticeable that in computer engineering department, freshmen are the much more likely to rate 5 than the others. On average, the gap[9] between a common CMPE freshmen and the other student levels range from -7% to 34.75% which means that out of 100 computer engineering students, there can be up to 35 more freshmen who would rate 5 although there is a chance that at most 7 less freshmen could rate 5.



The second interaction shows that the association between course level and rating depends on departments. The predicted probabilities clearly show that the course level has different effects on ratings according to the department. But the estimated probabilities in graduate courses have large variance due to small sample size except for CMPE department which makes it hard to distinguish the effect of departments on ratings between upper division and graduate courses[10]. A CMPE student is 9%~33% more likely to rate 5 in graduate course than upper division courses[11].

Figure 12. Predicted probability of rating 5 when a senior student takes a class of 25 people in Fall and expects 'A'

[9] 95% wald confidence interval of a common CMPE freshman's probability is (.48, .78), the intervals for sophomore, junior, senior, and graduate students are (.34, .54), (.44, .53), (.51, .56), and (.44, .57) respectively. The interval of gap is computed as $\left(.48-\frac{.54+.53+.56\quad .57}{4}, .78-\frac{.34\quad .44\quad .51+.44}{4}\right)$.

[10] For the 4 departments, 95% C.I.'s for upper division was nested inside C.I. for graduate division caused by small sample: While 3949 students took graduate CMPE courses, only 580 students took graduate courses in total in the other departments.

[11] The 95% wald confidence interval for graduate course is (65%, 84%) and for upper division course is (51%, 56%)

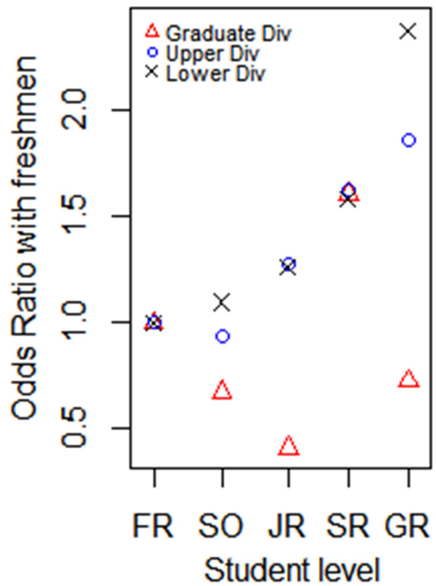


Figure 13. Odds ratio with Freshmen as reference level within each course level

The last interaction between student level and course level is shown in terms of odds ratio against freshmen in figure 13. In undergraduate courses, the academically maturer students are more likely to give rating 5[12]. 98% of the students who take graduate courses are graduate students, so the odds-ratios in the plot may not be reliable due to the difference of sample sizes. Among the rest of the 2%, the proportion of seniors getting ‘A’ is more or less the same as that of graduate students getting ‘A’ which could indicate that these extraordinary seniors can be considered as academically mature as graduate students. It is noteworthy that these seniors are more likely to be generous in rating than graduate students by the confidence interval of odds ratio (1.5, 3.1)[13] which means the odds of seniors submitting score 5 in graduate courses is about 1.5 to 3 times higher than the odds of graduate students. As a side note, there were only 4 juniors, 0 sophomore, and 2 freshmen who took graduate division courses, thus their estimated odds ratios were unreliable due to large variance.

Undergraduate students do not show particular difference in rating between lower and upper division courses. However, graduate students are quite different in rating by being the most generous in undergraduate courses while being the least generous in the course of their own level. Their odds of rating 5 in upper division courses is 2.3~2.8 times higher than the odds in graduate courses.

There were four outliers whose standardized deviance residuals were extremely large exceeding 10,000 in absolute value[14]. The outliers were from small departments that account for less than 0.6% of total students and their level in their own department was also very rare. Three of them expected grade ‘A’, but did not submit rating score 5 and the other one expected grade ‘B’, but submitted rating score 5. The estimated variance of the rating was very high with very small point estimates. One could consider removing these extreme outliers for more accurate predictions.

From the partial residual plot of class size[14], there seems to be a weak negative linear relationship between class size and rating which may satisfy one of the assumptions in the logistic regression model.

---

[12] Although sophomores seem less likely to be generous in upper division course than freshmen, the odds ratio is only 0.93 which is almost 1 thus the difference is negligible.

[13] Derivation in the appendix

[14] Index plot, Cook’s distance plot, Partial residual plot, Added-variable plots, and Constructed variable plot are in the appendix

For testing the added variability of the 6 main effects, added-variable plots for some indicator variables from main effect were shown in the appendix. As expected from figure 9, semester and course level seemed to show almost no effect as the straight line of LOWESS smoothing curve is shown. Student level has some weak association with rating. Being a senior adds a little more probability onto submitting rating score as 5.

The rest of the main effects, expected grade, departments, and class size showed relatively strong relationship with rating. For example, expecting a grade 'B' noticeably lowers the chance of rating score 5, majoring in computer engineering slightly raises the overall chance of rating score 5, and class size has weak negative association with rating.

The constructed variable plot of class size shows the approximately straight horizontal line indicating that the slope of standardized deviance residual against the constructed variable of class size is very close to zero which means no particular transformation of class size is necessary to satisfy the linearity assumption of the model.

To measure the predictive power of the model and find the optimal threshold for cutoff probability, a receiver operating characteristic(ROC) curve has been used[15]. AUC of the model is 0.71 which means if one's fitted probability is higher than another, s/he is 71% more likely to submit score 5. This model adds 21% more chance to the actual probability than a random classifier which correctly classify the ratings for the 50% of the time for every possible cutoff. The ROC curve of the model shows higher TPR than FPR for all possible cutoff probabilities which indicates performing better than the random classifier. The optimal cutoff probability of the model that maximizes TPR and minimizes FPR(which is equal to maximizing TPR – FPR, a.k.a. Youden's J statistic which corresponds to the vertical distance from the dashed line in the plot) can be found as 0.43. With this cutoff probability, the TPR is 0.69 and the TNR is 0.61. The predictive power of the model is not strong with low accuracy of point estimates. It may be necessary to investigate the inflated variance in coefficients that may be due to the outliers or insufficient number of main effects or interaction effects.

Summarizing the results of model adequecy checking, there are four extreme influential outliers which may change the implication of the effects in the model. In this report, however, they were included in every analysis. Although course level does not seem to have effects on the rating as a main effect in the model, there is noticeable interaction effect with department and student level, thus it may remain in the model to allow flexibility in the prediction curve. Semester, on the other hand, may be removed since it does not seem to affect the rating neither as a main effect nor as an interaction effect.

The teaching evaluation survey is not independent of student's characteristics including their expected grades, their tendancy to overestimate, and their academic levels. The strong association between their expected grades and ratings explains the inflation for both. Students

[15] The ROC curve is in the appendix.

have strong desire to achieve ‘A’ and this often leads them to be less objective about themselves and towards evaluating their instructors. Students with such desire may believe that getting an ‘A’ means that they digested most of the materials taught in class which makes them believe their instructor has very effective teaching skills. Considering that 40% of the grades are ‘A’, it may be unlikely that about 40% of the instructors have very effective teaching skills.

Rating also depends a lot on class size. Professors who teach in class with less than 40 students definitely have advantage over those who teach in larger classes and many departments from college of science have superlarge classes which lead to poor rating.

The size of departments vary so widely that the impact that one student carries also vary too much which may have unnecessary effect on the average rating as well.

Also, an instructor who teaches in both undergraduate and graduate courses may receive different ratings depending on the level of the students taking the class in a particular semester. Ratings in a graduate course with many seniors or an undergraduate course with many graduate students may be unusually high.

The teaching evaluation scores can be inevitably swayed by various aspects of students which may distort one’s teaching skills. Heavy reliance on the average score for promotion results in inflation of grades which makes students overlook the deeper sense of academic learning and instructors become less prone to challenge students with things that deepen their knowledge.

<Appendix>

#4. The model specification is in ‘SOTES_model_Jung-a Kim.pdf’ file with estimates to the nearest thousandth and origianal estimates

#5. Test of independence between rating and the main effects

Department vs Rating 5

```
Chi-squared test of independence at level 0.05
Pearson statistic = 1889.223
likelihood-ratio statistic = 1895.95
degrees of freedom = 46
critical value = 62.8296
```

Semester vs Rating 5

```
Chi-squared test of independence at level 0.05
Pearson statistic = 13.6712
likelihood-ratio statistic = 13.6675
degrees of freedom = 1
critical value = 3.8415
```

Course level vs Rating 5

```
Chi-squared test of independence at level 0.05
Pearson statistic = 288.843
likelihood-ratio statistic = 289.651
degrees of freedom = 2
critical value = 5.9915
```

Student level vs Rating 5

```
Chi-squared test of independence at level 0.05
Pearson statistic = 252.5054
likelihood-ratio statistic = 253.6293
degrees of freedom = 4
critical value = 9.487
```

Class size vs Rating 5

```
Chi-squared test of independence at level 0.05
Pearson statistic = 2459.431
likelihood-ratio statistic = 2586.729
degrees of freedom = 171
critical value = 202.5126
```

Expected grade vs Rating 5

```
Chi-squared test of independence at level 0.05
Pearson statistic = 7886.415
likelihood-ratio statistic = 8221.793
degrees of freedom = 3
critical value = 7.8147
```

#6-1. Positive association between course level and rating

	Rating 5	
course level	Yes	No
Lower	0.15	0.21
Upper	0.25	0.27
Graduate	0.06	0.06
$G^2 = 290,$	$\chi^2_{2,0.05;} = 6$	

Contingency table between course level and rating

	Rating 5	
course level	Yes	No
Upper	.28	.31
Lower	.17	.24
$G^2 = 281, OR = 1.29$		

Subtable to compare upper and lower division courses in rating

	Rating 5	
course level	Yes	No
Grad	.06	.06
Undergrad	.40	.48
$G^2 = 9, OR = 1.07$		

Subtable to compare graduate and undergraduate courses in rating

#6-2. Positive association between student level and rating

Student level	Rating 5: Yes	Rating 5: No
Freshman	0.06	0.09
Sophomore	0.06	0.08
Junior	0.10	0.12
Senior	0.17	0.19
Graduate	0.06	0.07
$G^2 = 254$,	$\chi^2_{0.05;4} = 9.5$	

Table between student level and Rating 5

student level	Rating 5: Yes	Rating 5: No
Soph.	.22	.29
Fresh.	.20	.30
$G^2 = 14, OR = 1.1$		

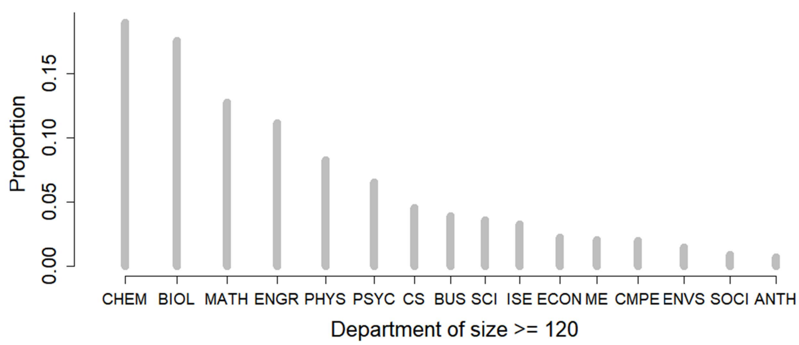
student level	Rating 5: Yes	Rating 5: No
Junior	.20	.24
Fresh+ Soph.	.23	.33
$G^2 = 91, OR = 1.2$		

student level	Rating 5: Yes	Rating 5: No
Senior	.20	.21
Fresh+Soph +Junior	.26	.33
$G^2 = 126, OR = 1.2$		

student level	Rating 5: Yes	Rating 5: No
Grad	.06	.07
Under-grad	.39	.48
$G^2 = 22, OR = 1.1$		

Subtables of proportions between student level and rating 5

#8. Proportions of departments in large classes



Proportion of departments of class size greater than 120

#13. Derivation for odds ratio wald confidence interval

Notation: gr = graduate student level, grd = graduate course level, rate 5 = submiting score 5

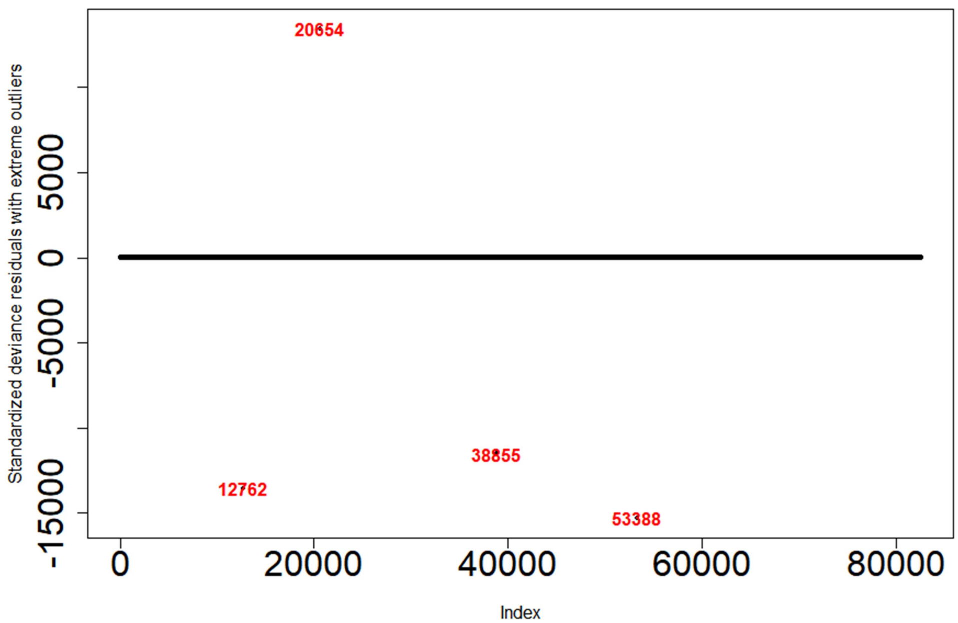
$$\log(\widehat{\mathrm{OR}}) = \mathrm{logit}(\hat{\theta}_{sr}) - \mathrm{logit}(\hat{\theta}_{gr}) = \hat{\beta}_{sr} + \hat{\beta}_{grd*sr} - (\hat{\beta}_{gr} + \hat{\beta}_{grd*gr}) \approx .7912$$

$$\hat{V}(\log(\widehat{\mathrm{OR}})) = \frac{1}{n_{gr,rate5}} + \frac{1}{n_{gr,not\ rate\ 5}} + \frac{1}{n_{sr,rate5}} + \frac{1}{n_{sr,not\ rate\ 5}} \approx .0327$$

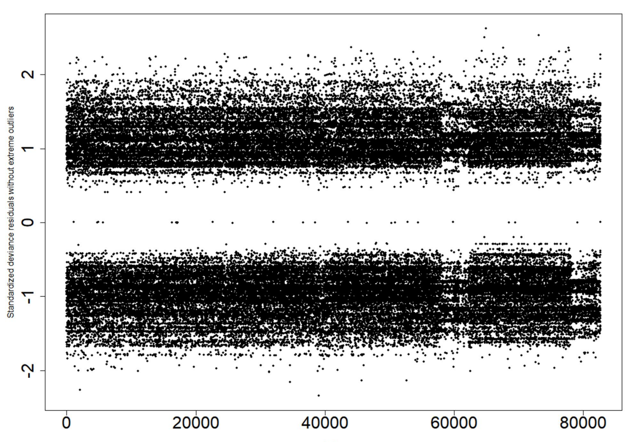
$$\mathrm{C.I.}^{95\%}_{\log(\mathrm{OR})} = \left(\log(\widehat{\mathrm{OR}}) \pm 1.96 * \sqrt{\hat{V}(\log(\widehat{\mathrm{OR}}))}\right) = (.4368, 1.1456)$$

$$\mathrm{C.I.}^{95\%}_{\mathrm{OR}} \approx (\exp\{.4368\}, \exp\{1.1456\}) \approx (1.55, 3.14)$$

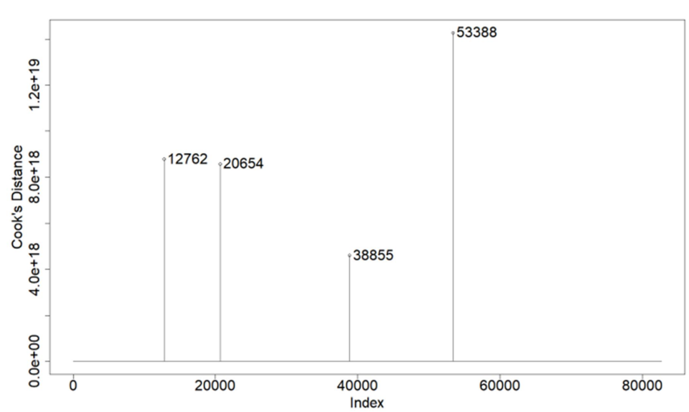
#14-1. Index plot of standardized deviance residuals with the four extreme outliers



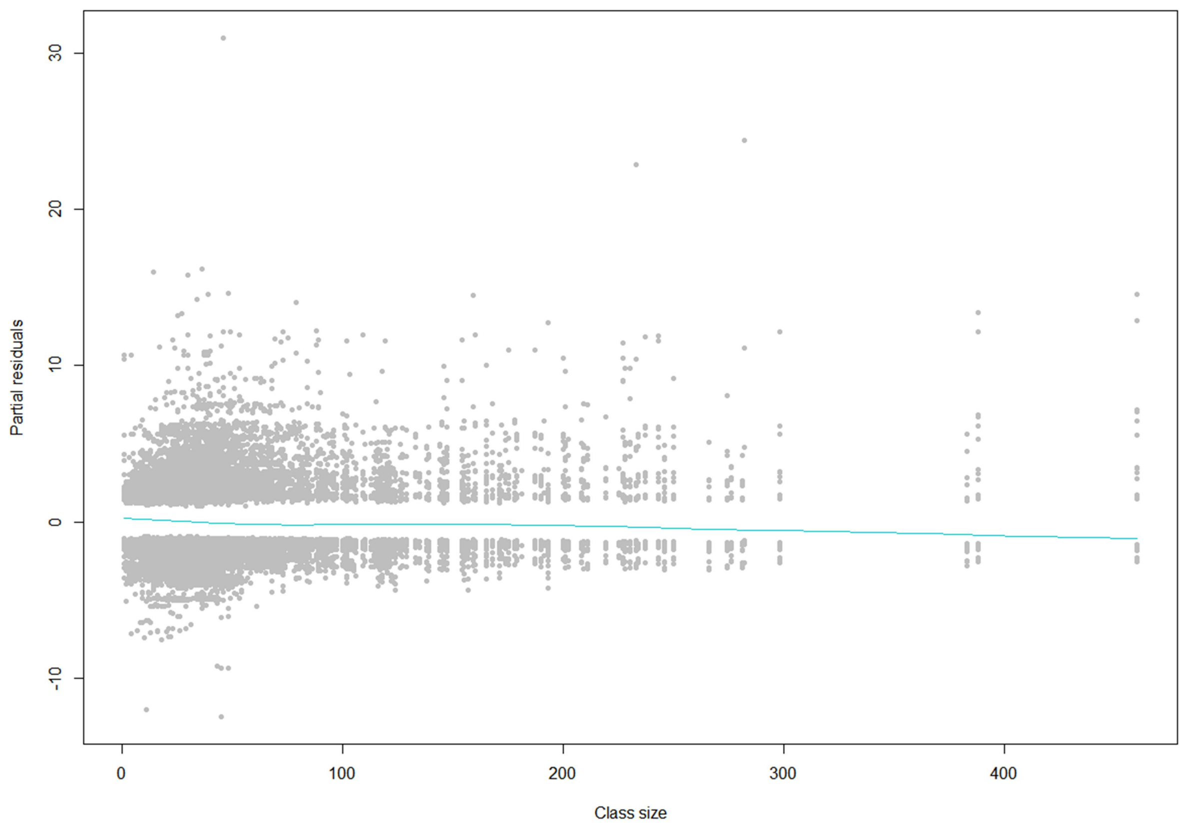
#14-2. Index plot of standardized deviance residuals without the four extreme outliers



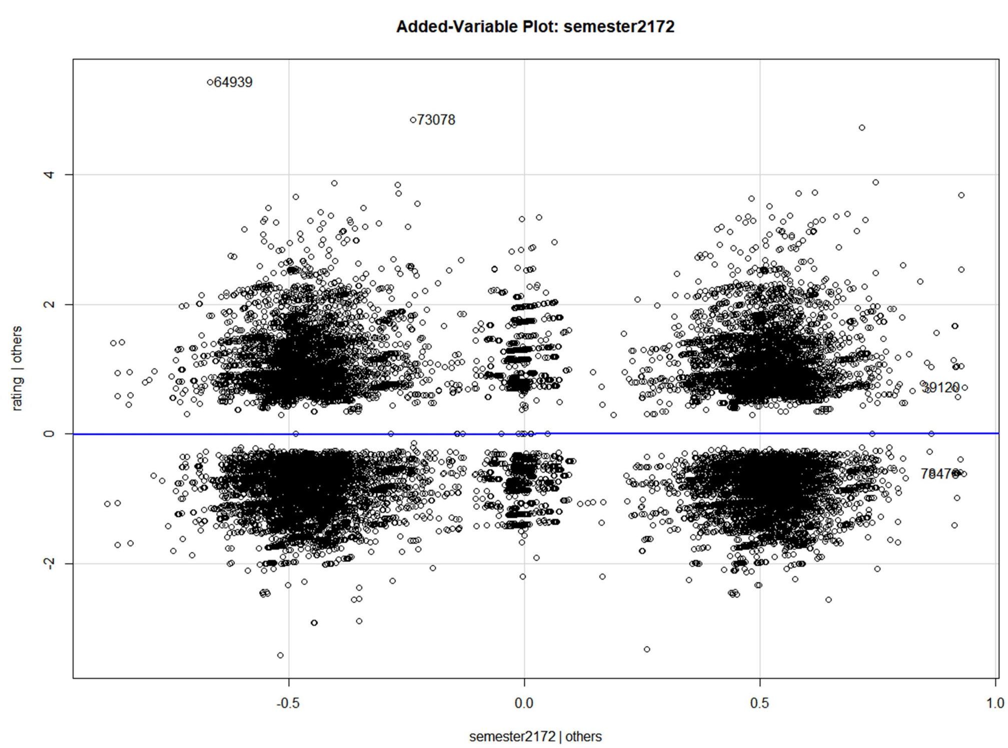
#14-3. Cook's distance plot



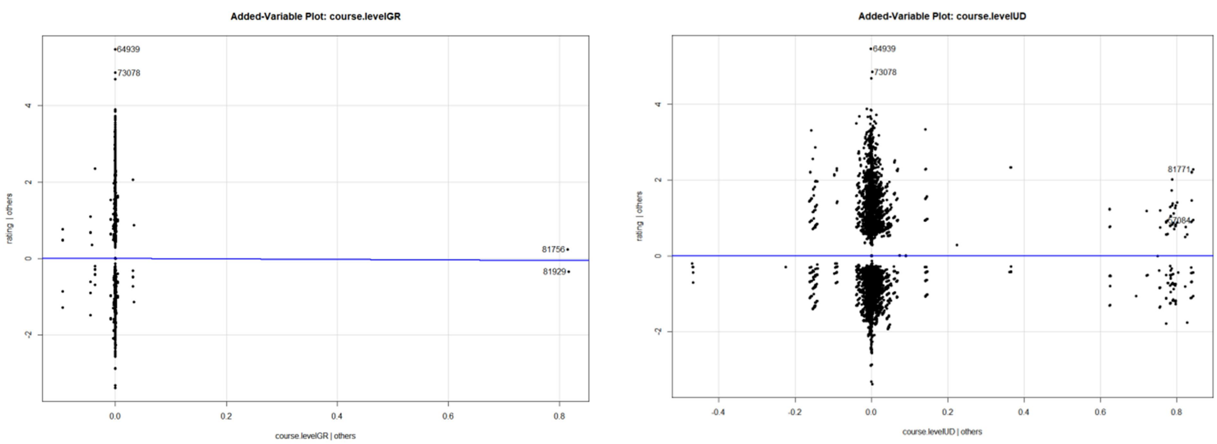
#14-4. Partial residual plot of class size



#14-5. Added variable plot of Spring semester showing the straight line of LOWESS curve

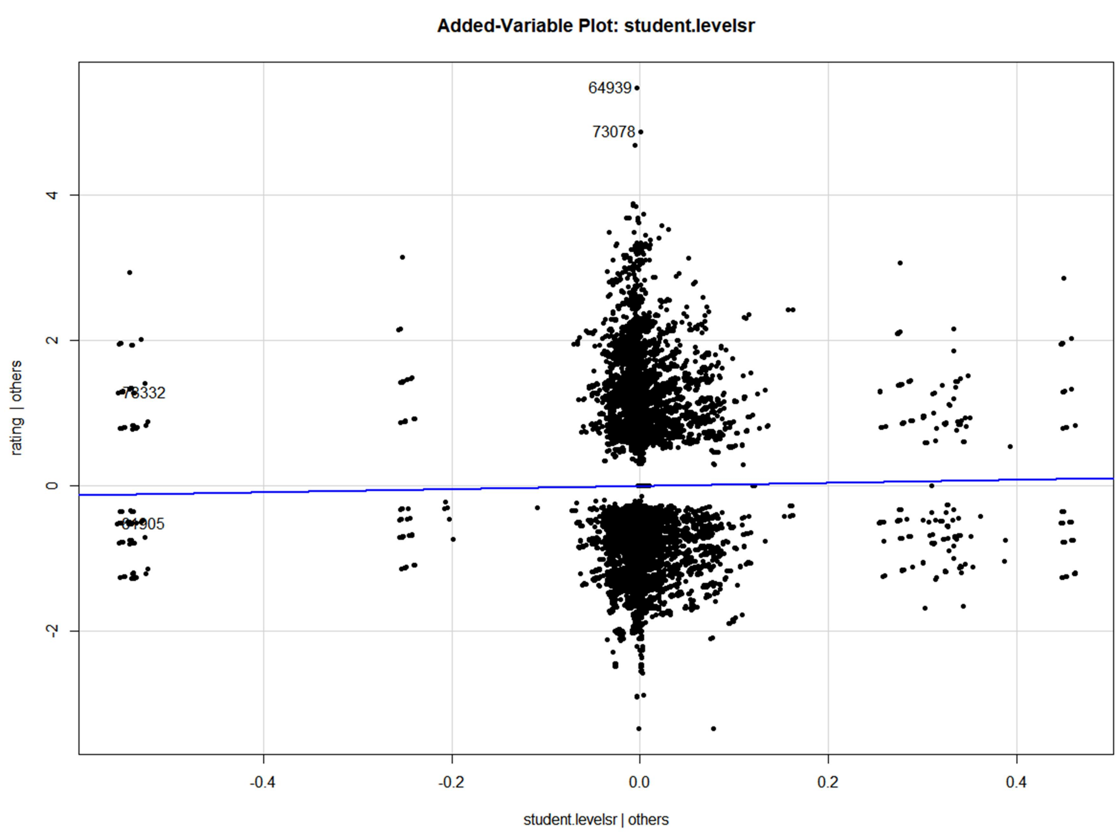


#14-6. Added variable plot of Graduate course(left) and upper division course levels(right) showing the straight line of LOWESS curve

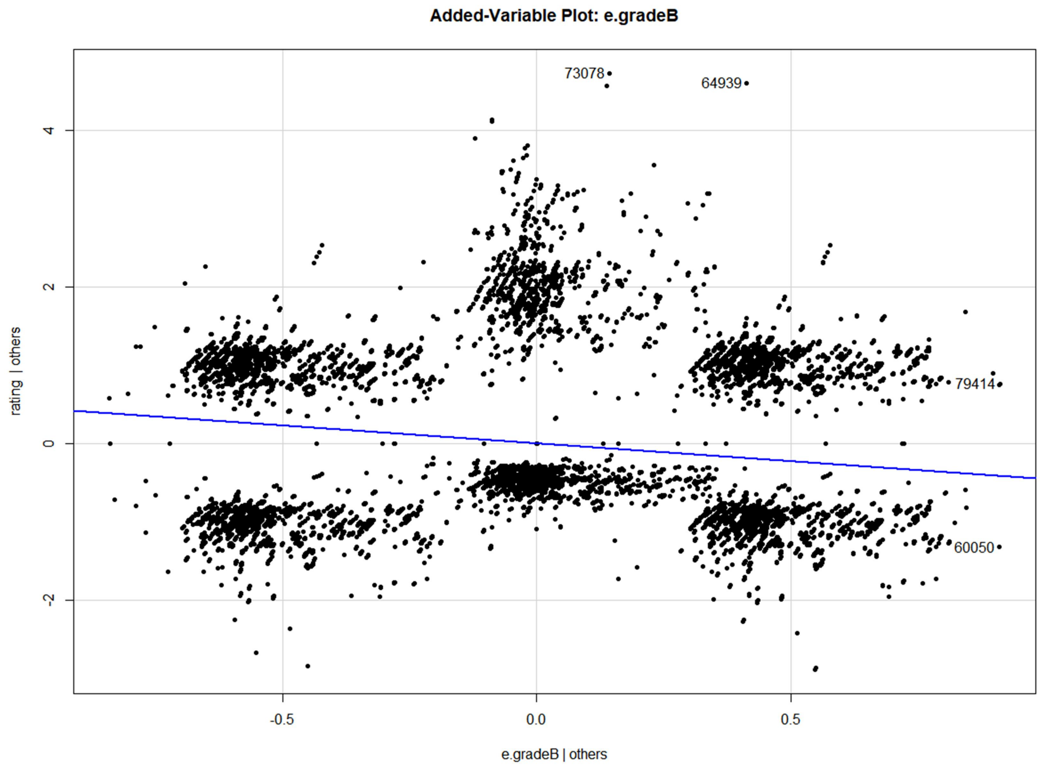


* The slight change of slope in LOWESS smoothing line in the left plot is only due to the two observations with high leverage

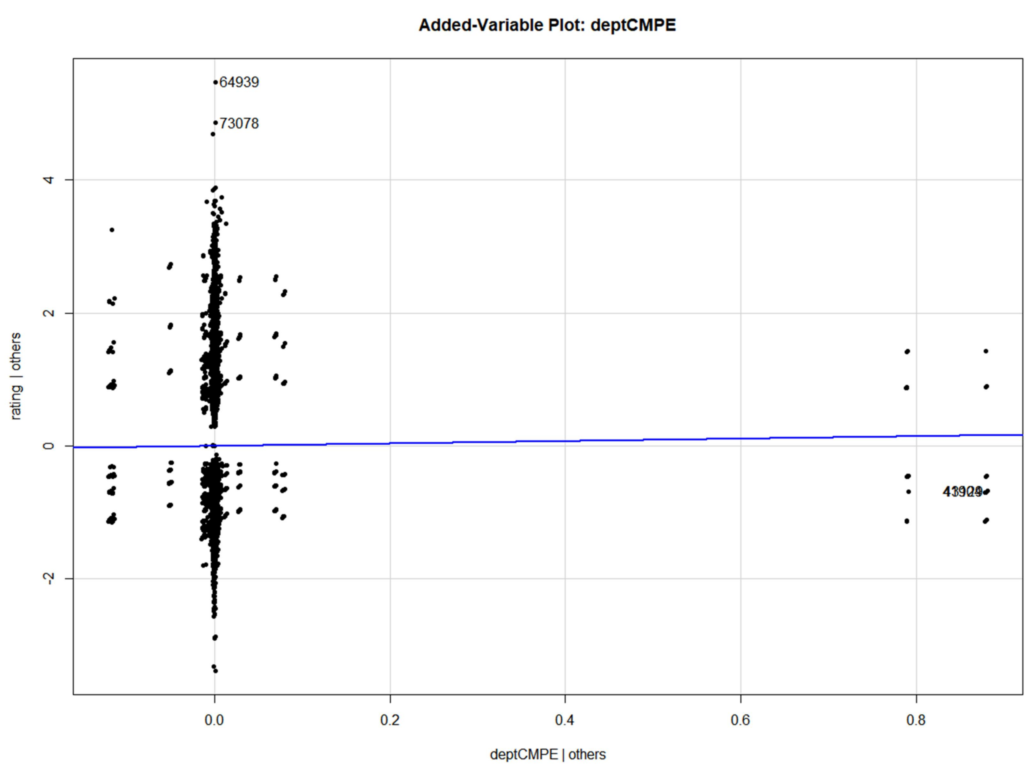
#14-7. Added variable plot of student level(senior)



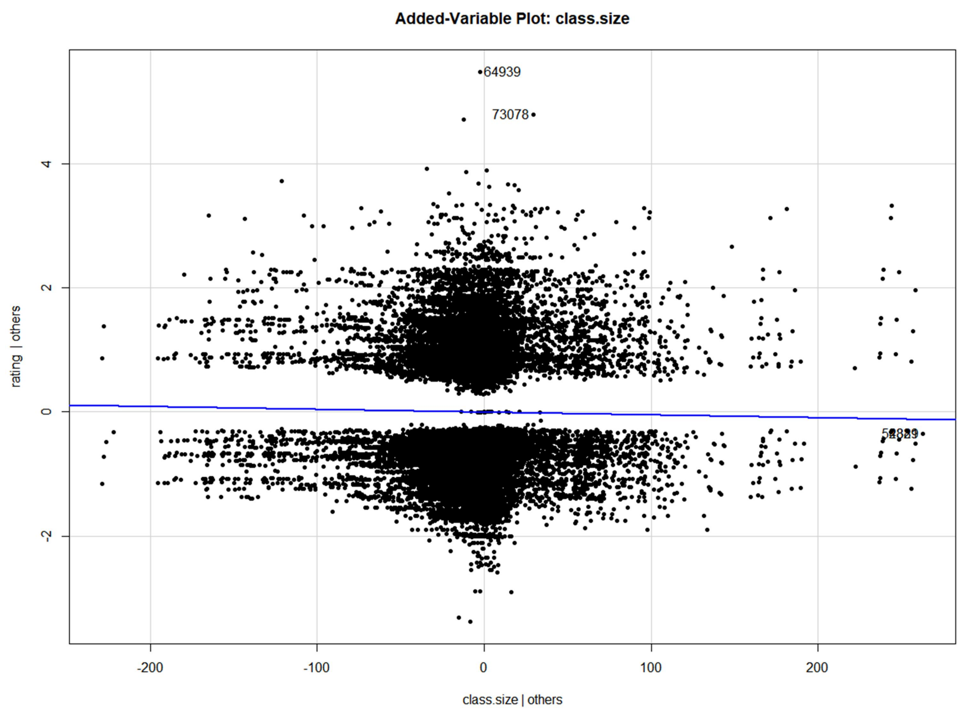
#14-8. Added variable plot of expected grade as 'B'



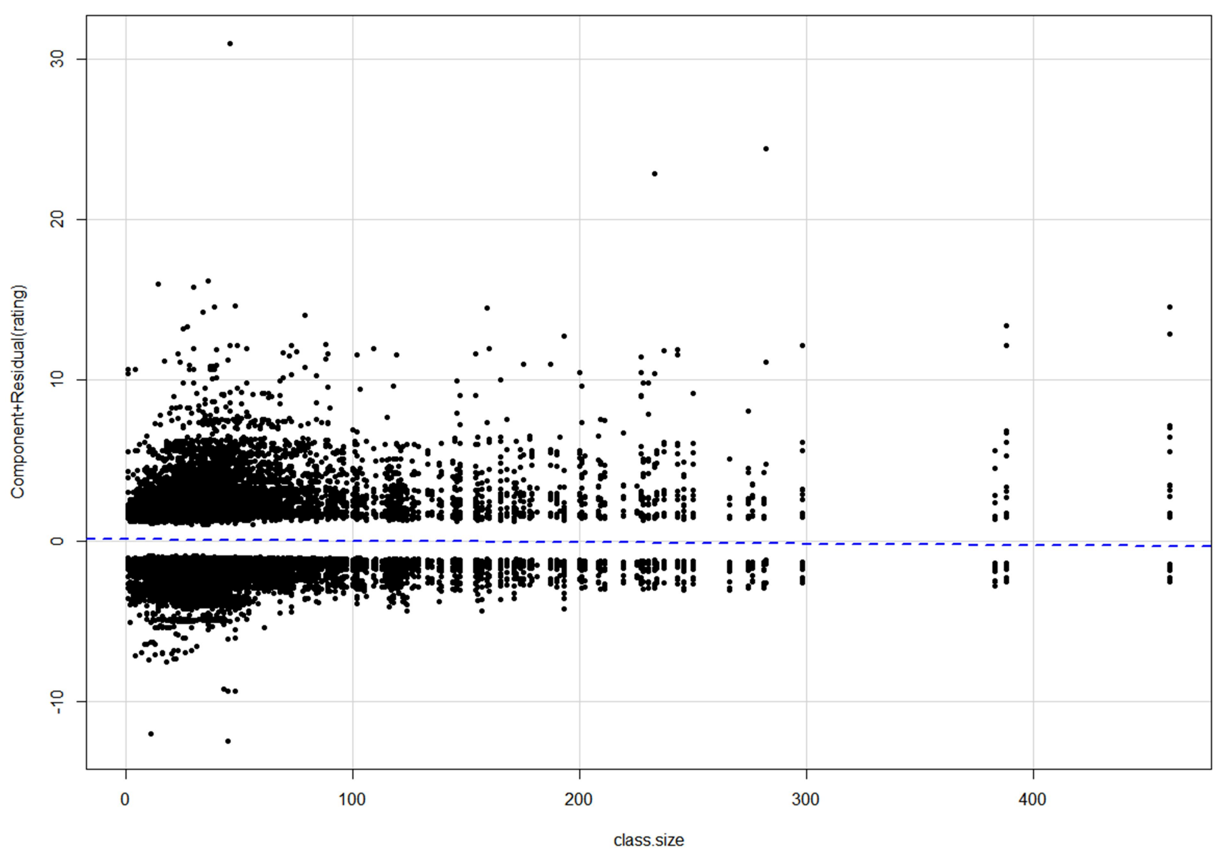
#14-9. Added variable plot of Computer Engineering department



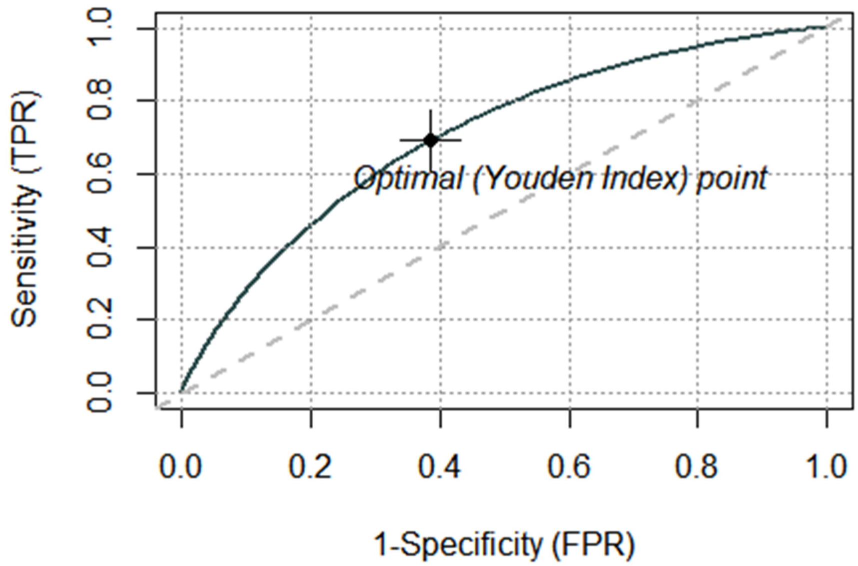
#14-10. Added variable plot of class size



#14-11. Constructed variable plot of class size



#15. ROC curve of the model



< References >

Lee, B. (2019). MATH258: Categorical Data Analysis, topic A-F, [PDF slides].

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Andri Signorell et mult. al. (2019). DescTools: Tools for descriptive statistics. R package version 0.99.29.

Youden's J statistic. (n.d.). In *Wikipedia*. Retrieved December 17, 2019. from https://en.wikipedia.org/wiki/Youden%27s_J_statistic