ECON 4848-004
Haley Miller, James Paterson, and Anna Sanders
Dr. Kaffine
Fall 2019

Average Grade's impact on Overall FCQ Score

I. Introduction

A. Significance

At the University of Colorado Boulder, all students have the opportunity to submit Faculty Course Questionnaires, or more colloquially FCQs, to evaluate each class and professor at the end of each semester. The University of Colorado Boulder asks a combination of multiple choice and short answer responses in order to assess students' overall learning experience. These results are important not only to the university's own assessment of classes and professors, the data is also helpful to students weighing the options of which classes to take and which professors to take them from. Similarly, studying trends within this data set could better help the university adapt classes to be better in terms of teaching, content, and fairness.

B. Research Question

Our overall research question is how student's grades impact Overall Class FCQ Scores. In order to further understand the impact grades (or student's perceived grade as FCQs are submitted before average class GPA is) on overall class rating, we have split this question into multiple parts, holding different variables constant, and finally creating one comprehensive regression to answer this question.

II. Data

A. Data Set

The University of Colorado Boulder makes the multiple choice section of the FCQ public through an online database and a master Excel sheet. In this data set, we used the Excel sheet's Data section as it had a complete table of each course, along with a list of variables associated with each class, including overall grade, hours, department, and more. The FCQ Excel sheet is updated each semester from Fall 2006 to the most current term year, which right now is Spring 2018. For research purposes, we decided to omit all data pertaining to graduate level courses, as different departments have different requirements for enrollment amount and grade distributions. We felt that graduate classes could skew the data as graduate classes are usually smaller and taught only by university professors and many graduate programs have hard caps of GPA causing the average grade for graduate level classes to be inflated.

B. Important Features

The data features 50 variables relating to each class taught from Fall 2006 to Spring 2018. Variables included those for average FCQ score, average grade, department, college,

subject, course number, section number, class size, credit hours, percent A, B, C, D/F, instructor name, workload, instructor average, and instructor type.

Our two main variables, AVG_GRD and AvgCourse are continuous variables. AVG_GRD is continuous from 1-4 (alphabetical grades 1=F, 2=C, 3=B, and 4=A) while AvgCourse is continuous from 1-6 (with 6 being the highest possible rating).

Since the data taken from the FCQ results was already averaged over the form responses received, we did not know the true distribution of scores or if there were any outlier responses per class; however, we still felt that this data was still an accurate representation for each variable reported.

In our analysis, we mainly used the two variables for overall grade (AVG_GRD) and overall FCQ class rating (AvgCourse). In order to better answer our research question, we used the variables of enrollment (N_ENROLL), credit hours (Hours), instructor type (adapted from insttl1), FCQ's returned (N_Ret), term year (num_year), department (Subject), and upper/lower division (Level), average instructor FCQ score (AvgInstructor), and workload (Workload_Raw), to hold overall class rating and overall grade constant against.

III. Methodology

A. Data Cleaning Variable Creation

As stated in the previous data section, we chose to omit data from graduate level programs (classes with course numbers greater than or equal to 5000). To do this, we used the code "keep if" to keep only upper and lower division classes. We similarly checked for missing or unreported average grades and average FCQ scores using the same command.

In order to run certain summary statistics and regressions, we created one new variable and three new group variable. When looking at the instructor level variable closer (insttl1), we found that there were over 100 listed instructor levels. Instead of running a regression with these levels, we decided to sort the levels into three different categories: one for grad students, one for all instructors, and one for professors and assistant professors. We used the replace command on our new named variable instlvl to assign insttl1 titles to numbers (1 for grad, 2 for instructors, and 3 for professors). The group variables were generated through the code "egen newvar= group(oldvar)" for YearTerm, Subject, and Level, which are all categorical variables.

B. General Statistics

To begin answering our research question, we first created tables of summary statistics that could help answer our overall question and our sub-questions holding different variables in the data constant. From these tables, we were able to get a sense of the dataset and important features of the data set.

We also created multiple scatter plots displaying the relationship between average grade and average course score and comparing against the relationship between average grade and average course score of certain observations, as well as the time series relationships of average course score and average grade over time. The results of both of these findings are further detailed in the findings section.

From these initial findings, we decided on nine variables we wanted held constant in a regression of AVG_GRD against AvgCourse. All results are rounded to 2 significant figures.

C. Regressions

We felt like there was no need to apply a logarithmic scale to any of the variables as the numbers themselves represent real results (i.e. a one grade level increase (C to B) will lead to a change in average FCQ rating).

Our first regression was the simple model regressing AVG_GRD against AvgCourse, holding no variables constant. This regression gave us a baseline for the AVG_GRD coefficient to compare all other regressions against. The equation for this regression is $y = a + bx$, where $y$is estimated overall FCQ score, a is the y-intercept (_cons coefficient in the regressions) and $b$is the marginal effect of grades on FCQ score (AVG_GRD coefficient in the regressions).

Each successive regression regressed overall grade against overall class FCQ score holding one categorical variable constant. We decided that this was the most effective way to find the relationship between the single variable held constant with our x and y variables. The equation for these regressions is $y = a + b1x1 + bn * xn$ where $a$is the y-intercept (_cons coefficient in the regression), $b1$is the marginal effect of grades on FCQ score, $x1$is the x-variable for grades, bn is the coefficient for the categorical variable (represented by numbers (n) corresponding to the correct category), and $xn$is the x-variable corresponding to the category (0 to not include, 1 to include). Since the categorical coefficients reported were one less than the actual amount of categories, we can interpret the regression $y = a + b1x1$where $x2 = 0$as the regression holding everything except for the first category constant.s

Our data was both categorical and time series, so we used the regression techniques for panel data. For each of our categorical data, we regressed with the extra term i.[xvar], regressing only against our x and y variables (AVG_GRD and AvgCourse).

Our general interpretation for these results was that the higher the coefficient on the AVG_GRD variable, the higher the effect grades (or perceived grades as the FCQ form must be filled out before grades are released) has on class rating. This could mean that the students were less happy with the course as a whole as class rating was highly dependent on course grade. All results for the AVG_GRD coefficients were taken as economically significant as the AvgCourse variable had such a small bound (1-6). Similarly, statistical significance was taken at p-values below 0.05, or 95% confidence.

IV. Results

A. Summary Statistics

a. General Summary Findings

Our summary statistics found that overall, the University of Colorado: Boulder's average GPA was 3.17 (about a B average), average course FCQ rating was 4.78 (out of 6), average number of students enrolled per class was 40.64 people, and average credit hours of a class was 2.9. When sorting by year, both the average grade and average course FCQ rating has gone up

from 3.1 to 3.2 and 4.66 to 4.8 respectively. When looking at the averages per department, we found that different departments have wildly different averages for grades and course FCQ ratings. Looking specifically at the ECON Department, the average grade is 2.73 (about a B-) while the average course is a 4.51, which are both lower than the University average. The ECON Department has the third lowest average grades (only APPM and MATH's average of 2.71 beats this number). Looking at a department at the other end of the spectrum is EDUC; with an average grade of 3.72, it has one of the highest average grades, however, its average FCQ score is only 4.90.

We also found that upper division classes, on average, had higher average FCQ ratings and higher average grades (4.71 to 4.85 and 3.10 to 3.26 respectively). Similarly, across instructor types (comparing grad students, instructors, and professors), average FCQ scores and average grades were mostly the same with very small amounts of variation in the averages. Larger enrollment also led to lower average FCQ scores and average grades.

b. Scatter Plot Findings [insert the scatter plots below]

To do some general analysis on the relationship between grades and FCQ score, we made a scatter plot and graphed the linear relationship based on just these two variables.

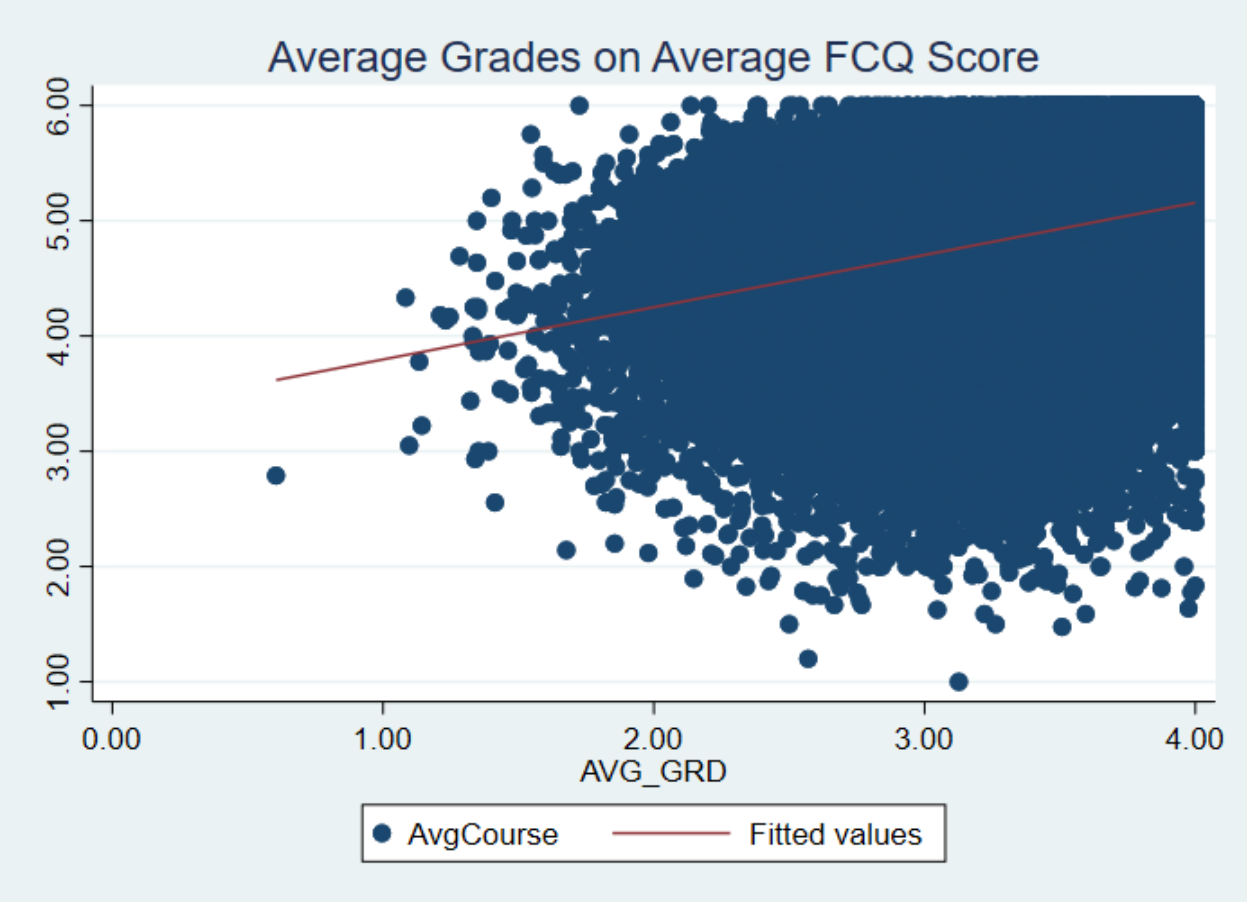

Figure 1 General Scatter Plot from code scatter AvgCourse AVG_GRD || lfit AvgCourse AVG_GRD, title("Average Grades on Average FCQ Score").

As you can see above, the scatter plot is dense around the right most quadrants as most classes at the University fall within the rage of grades > 2.5 and FCQ score > 4. The data, however, still looks to follow a linear relationship more so than a quadratic or other non-linear relationship. The scatter plot models this basic linear relationship with the red line labeled "Fitted values".

B. Regression Results

a. General Regression

The results of the general regression just for average grade and average FCQ score is below. All subsequent regressions were not included, but the equations were taken directly from result tables like the one below.

```
. * General Regression
. regress AvgCourse AVG_GRD, robust

Linear regression                               Number of obs     =     67,884
                                                F(1, 67882)       =    6803.81
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0920
                                                Root MSE          =     .63186

------------------------------------------------------------------------------
             |               Robust
   AvgCourse |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     AVG_GRD |   .4532947   .0054955    82.49   0.000     .4425236    .4640658
       _cons |   3.343462   .0175996   189.97   0.000     3.308967    3.377957
------------------------------------------------------------------------------
```

Figure 2 General Regression results from code regress AvgCourse AVG_GRD, robust.

The general equation for the relationship between average grade and average FCQ score is: $y = 0.45x + 3.34$ where y is the average FCQ score and x is the average grade for the class. This result is statistically significant as the p-value is less than 0.05. We interpreted this regression as a one letter grade increase in average grade corresponds to a half a point higher FCQ score. Of course, this regression takes into account nothing other than average grade, so below, we have held constant other variables we thought could be significant.

b. Holding Class Size Constant

As the University of Colorado Boulder is a large school, there tend to be some fairly large classes (upwards of 300 students). Running this regression would allow us to see if the amount of students enrolled greatly impacted average FCQ scores.

The general equation for the relationship between average grade and average FCQ score is: $y = 3.52 + 0.41x - 0.001z$ where y is the average FCQ score and x is the average grade for the class and z is the number of students enrolled. This result is statistically significant as the p-value is less than 0.05. We interpreted this regression as a one letter grade increase in average

grade corresponds to a 0.41 point change in FCQ score and a 1 person increase in enrollment leads to a -0.001 point change in FCQ Score.

c. Holding Credit Hours Constant

The university also has classes with different credit hours, which roughly translates to how many hours spent in class. We wanted to make sure that we held credit hours constant as it is possible that credit hours had a large effect on overall FCQ score.

The general equation for the relationship between average grade and average FCQ score is: $y = 2.90 + 0.50x + 0.10z$ where y is the average FCQ score and x is the average grade for the class and z is the credit hours of the class. This result is statistically significant as the p-value is less than 0.05. We interpreted this regression as a one letter grade increase in average grade corresponds to a 0.50 point change in FCQ score and a 1 credit hour increase leads to a 0.10 point change in FCQ Score.

d. Holding Instructor Type Constant

As the University of Colorado Boulder is a large school, there is a large faculty with many different types of lecturers/instructors. Running this regression would allow us to see if the type of teacher (grad student, instructor, or professor) greatly impacted average FCQ scores.

The general equation for the relationship between average grade and average FCQ score is: $y = 3.32 + 0.45x + 0.01z_2 + 0.04z_3$ where y is the average FCQ score, x is the average grade for the class, and z_2 and z_3 are binary terms representing if the instructor is an instructor or professor respectively. This result is mostly statistically significant as the p-value is less than 0.05 for all terms exceptz_2. We interpreted this regression as a one letter grade increase in average grade corresponds to 0.45 change in FCQ score and a change in instructor type from graduate student leads to an either a 0.01 or 0.04 change in FCQ Score.

e. Holding FCQs Returned Constant

The FCQ form is not mandatory, so depending on the class, amount of FCQ responses can widely vary. If only a few FCQs are returned, the observed average FCQ score could be inflated or deflated depending on who submitted the form. We decided to hold this variable constant in addition to number enrolled to ensure higher or lower FCQ forms returned did not create a biased estimate.

The general equation for the relationship between average grade and average FCQ score is: $y = 4.47 + 0.43x - 0.001$ where y is the average FCQ score, x is the average grade for the class, and z is the number of FCQ forms returned. This result is statistically significant as the p-value is less than 0.05. We interpreted this regression as a one letter grade increase in average grade corresponds to 0.43 change in FCQ score and an increase of 1 FCQ form returned corresponds to a -0.001 change in FCQ score.

Feedback from the in-class presentation gave us feedback to compare the percent returned in addition to the number returned. Holding the percent FCQ return rate may give different answers as classes at the university have a wide variety of enrollment, so 20 returned FCQ forms would have a different overall effect if the class size was 25 rather than 300.

The general equation for the relationship between average grade and average FCQ score is: $y = 3.14 + 0.43x + 0.39z$ where y is the average FCQ score and x is the average grade for the class and z is the percent of FCQs returned. This result is statistically significant as the p-value is less than 0.05. We interpreted this regression as a one letter grade increase in average grade corresponds to a 0.43 change in FCQ score and an increase of 1% FCQ forms returned corresponds to a 0.39 change in FCQ score.

f. Holding Term Year Constant

In our initial summary of the data, the results that the overall average grades and overall FCQ score for the University have been slowly increasing. As this data was taken from as early as Fall 2006, it was important to hold the year term constant as there was likely a significant effect on average FCQ score.

The general equation for the relationship between average grade and average FCQ score is: $y = 3.26 + 0.45x + z_i z$ where $y$is the average FCQ score, x is the average grade for the class, and z is a binary variable if it is the year, also relating to the year coefficient z_i. As there were above 20 year coefficients, we found it was easier to write the overall equation as such. This result is statistically significant as the p-value is less than 0.05 for AVG_GRD. As the year term variable was changed into a group variable, we can interpret larger numbers as higher years (e.g. Fall 2006 to Spring 2007 is a 1 unit increase). We interpreted this regression as a one letter grade increase in average grade corresponds to a 0.45 change in FCQ score and a per-semester increase leads to between a 0.02 and 0.17 change in FCQ score.

g. Holding Department Constant

Since the University of Colorado Boulder has many different departments, for this question, we decided to focus our analysis in the Economics Department, as we are all Economics Majors. As found above in our general analysis, the overall ECON average grades are much lower than other departments', making the results of the regression particularly interesting.

The general equation for the relationship between average grade and average FCQ score is: $y = 4.02 + 0.44x + z_i z$ where y is the average FCQ score, x is the average grade for the class, and z is a binary variable if it is that department, also relating to the department coefficient z_i. This result is statistically significant as the p-value is less than 0.05 for AVG_GRD. We interpreted this regression as a one letter grade increase in average grade corresponds to 0.44 change in FCQ score. As the Economics Department is coefficient (51), the equation for the relationship of average grades on average FCQ score in the Economics Department is $y = 4.02 + 0.44x - 0.15$.

h. Holding Course Level Constant

The University of Colorado Boulder also differentiates between Upper and Lower Division classes. Lower Division classes are usually very large and required or prerequisites for each department while Upper Division classes are smaller and cover topics within a certain subject. As one of our general summary tables found a large difference between both the average grade and average FCQ score, we felt it was important to hold this variable constant.

The general equation for the relationship between average grade and average FCQ score is: $y = 3.35 + 0.44x + 0.07z$ where y is the average FCQ score, x is the average grade for the class, and z is the binary variable if the class is an Upper Division Class (and not a Lower Division class). This result is statistically significant as the p-value is less than 0.05. We interpreted this regression as a one letter grade increase in average grade corresponds to a 0.44 change in FCQ score for Lower Division classes and an additional 0.07 change in average FCQ score if the class was Upper Division.

i. Holding Average Instructor FCQ Score Constant

As the average FCQ score is interpreted more as the class overall as a grade while average instructor FCQ score is only based on the instructor's performance, there is likely to be a large interaction effect as sometimes instructors can increase or decrease the class as a whole.

The general equation for the relationship between average grade and average FCQ score is: $y = 0.06 + 0.13x + 0.84z$ where y is the average FCQ score, x is the average grade for the class, and z is the average instructor FCQ score. This result is statistically significant as the p-value is less than 0.05. We interpreted this regression as a one letter grade increase in average grade corresponds to a 0.13 change in FCQ score and a one point increase in average instructor FCQ scores led to a 0.84 change in average FCQ scores. As these results seemed highly correlated with each other, we decided to run an interaction effect on average FCQ scores and average grades.

```
. regress AvgCourse AVG_GRD AvgInstructor c.AvgCourse#c.AvgInstructor, robust

Linear regression                               Number of obs     =     67,878
                                                F(3, 67874)       >   99999.00
                                                Prob > F          =     0.0000
                                                R-squared         =     0.9742
                                                Root MSE          =     .10655

------------------------------------------------------------------------------
             |               Robust
   AvgCourse |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     AVG_GRD |  -.0176727    .001078   -16.39   0.000    -.0197857   -.0155598
AvgInstructor|  -.5757346   .0048601  -118.46   0.000    -.5852604   -.5662087
             |
c.AvgCourse#c.AvgInstructor|   .1688165   .0004203   401.65   0.000     .1679927    .1696403
             |
       _cons |   3.582717   .0155919   229.78   0.000     3.552157    3.613277
------------------------------------------------------------------------------
```

Figure 3 Interaction Effect Regression from code regress AvgCourse AVG_GRD AvgInstructor c.AvgCourse#c.AvgInstructor, robust.

As the interaction term was a non-zero number and statistically significant and the individual term coefficients were negative and significant, we decided to leave average instructor FCQ score out of the final equation, as we were still unsure the effect average instructor FCQ had on average course, average grade, and other variables not in this regression.

j. Holding Average Work Hours Constant

The FCQ form also has a question on how many hours you worked on this class per week, including class sessions. This data is particularly relevant as some professors and subjects tend to have more out of class work (and thus higher average work hours per week). Reported weekly workload hours is reported in ranges (0-3, 4-6, 7-9, 10-12, 13-15, and 16+). As the Workload_Raw increased by one unit, the reported weekly workload changed by around 2 hours.

The general equation for the relationship between average grade and average FCQ score is: $y = 2.94 + 0.48x + 013z$ where y is the average FCQ score, x is the average grade for the class and z is the change average reported workload hours divided by two. This result is statistically significant as the p-value is less than 0.05. We interpreted this regression as a one letter grade increase in average grade corresponds to a 0.048 change in FCQ score and a one point increase (or around 2 hour increase) in raw work hours led to a 0.13 change in average FCQ score.

k. Final Regression Holding All Significant Variables Constant

From our results above, we found that [all variables tested] were statistically significant to average FCQ score. Thus, our final regression includes all variables except for average instructor FCQ.

As the entire equation would likely be too large to write out completely, we can interpret just the coefficient of AVG_GRD to find the effect of average grades on average FCQ score holding all else constant. The coefficient for average grade is 0.43, so we can interpret this as holding all else equal, a 1 letter grade increase of average grade leads to a 0.43 change in average FCQ score.

V. Conclusion

A. General Conclusion

From our analysis, we found that there is a statistically significant relationship between average grades and average FCQ score. Our research question of how students' grades affect FCQ score can be answered with the equation found in the final regression. This conclusion leads us to believe that the rating system could be biased such that higher grades directly correlate to higher FCQ scores.

B. Next Steps

The results found in this research paper should only be taken as preliminary findings as there are some major problems with this data set and interpretation. Firstly, this data may not be the most complete set of observations for every class. As FCQs are not mandatory, but highly encouraged, the amount of forms returned can sometimes be very low. This can skew the ratings to be higher or lower depending on those who submitted the forms. As overall grade is calculated from the entire class population, there could be a large discrepancy between the true average FCQ score and the observed FCQ results. Similarly, this paper was limited to only looking at the effect of average grades on average FCQ. Other variables, like instructor FCQ scores and

distribution of grades, could also be relevant to analyzing this data set; however, we were not able to run any regression or analysis of these variables.

Our research could also be expanded to better find the exact effect grades have on average FCQ scores. Due to time limitations, we were not able to run regressions to find the interaction effects of certain variables, which could better quantify the relationship between the variables in the data set. We especially would have liked to better quantify the relation between average FCQ scores and average instructor FCQ scores. Similarly, as we mainly focused on the Economics Department in our research, other departments may have higher or lower coefficients of average grades effect on average FCQ score.