Here is the Title of Your Paper

Alan Yeung[*]

02/24/16

Abstract

Here comes your abstract.

This is a template for documents prepared with LaTeX for the journal Economics-The Open Access, Open Assessment E-Journal (www.economics-ejournal.org).

You find further instructions in the Introduction. Please read them.

Keywords: Insert the keywords here, separated by commas.

Journal of Economic Literature Classification: Insert the classification numbers, such as J7, B54, etc. here.

[*]Give your affiliation and thanks here.

Contents

1 Introduction

In higher education, economists and prospective students often use the performance measures of the student body as a proxy measure for the quality of the institution (Smith and Stange, 2015; Sarmiento Espinel et al., 2015; Black and Smith, 2006). For example, the average high school GPA, average SAT score, and average class rank are all common reported characteristics of an institution and indicators of institutional quality. The underlying idea of using the quality of the student body as a measure for the quality of the institution, is that the quality of one's peers matter. Indeed, several studies have shown that peer quality has a positive impact on long term student outcomes, such as graduation rates and cumulative GPA (Smith and Stange, 2015; Luppino and Sander, 2015; Ost, 2010). Naturally, these findings lead to another area of research, identifying the specific areas of a student's life where peer effects are most significant. One of the most obvious areas for a strong effect to exist is in the classroom. Thus, we seek to understand peer effects in the classroom, specifically how the ability of one's classmates impacts a student's class grade.

This study contributes to the literature in two ways. First, this is the first study of it's kind (known to us) to research peer effects at the classroom level at this type of institution (a small private liberal arts college). Whereas the majority of other studies focus on K-12 education, we analyze the peer effects at an institution of higher education. Second, due to our access to data, we are able to use a unique method (at least in terms of classroom peer effects literature[1]) to analyze the data and combat selection bias, a serious problem in studies of this type (Carman and Zhang, 2012; Burke and Sass, 2013; Ding and Lehrer, 2007).

Using data from a small private liberal arts college in the mid-west we find that, regardless of the changes we applied to the peer measure, the proportion of high achieving students in a class has a small but significant negative impact on the grades of middle

[1] Most of the literautre uses quasi-random assignments or fixed effects models. See section 2 (Literature Review) for more details.

achieving students. Additionally, we find that, when the peer measure is at its strictest level, the proportion of low achieving students in a class has a small but positive impact on the grades of middle achieving students. These results are contrary to those found in most of the current literature (Kang, 2007; Carman and Zhang, 2012; Burke and Sass, 2013; Schlosser et al., 2008; Lavy et al., 2012), which may be due to changing peer dynamics from K-12 education to higher education, grade curving occurring in classes at the institution, or a combination of both. Though these results are interesting, it is unclear how well they generalize to other institutions, of different types and the same type, due to the current scarcity of literature.

The structure of the paper is as follows: We discuss the related literature in section 2. In section 3, we examine the data used in this study. In section 4, we describe our empirical methodology. We present our results, discuss the implications, and offer suggestions for further research in section 5. Lastly, we conclude in section 6.

2 Literature Review

Smith and Stange (2015), Luppino and Sander (2015), and Ost (2010) show that peer effects have important long term impacts on cumulative GPA and graduation rates at institutions of higher education. These are important broad findings, which have naturally led to a more narrow focus in the literature. Much of the current literature focuses on identifying the areas of a student's life where peer effects are most significant.

For instance, roommate peer effects research has been a trending topic, due to the natural quasi-random roommate assignments made by many institutions of higher education. Quasi-random roommate assignments, such as those used in Griffith and Rask (2014) and Zimmerman (2003), allow researchers to avoid the perils of selection bias and lead to theoretically more accurate results. However, most of the literature is divided on whether or not roommate peer effects exist (Griffith and Rask, 2014; Zimmerman, 2003; Sacerdote, 2000; Foster, 2006; McEwan and Soderberg, 2006). This is a testament to the difficulty of finding peer effects and the difficulty of solving the econometric problems that exist in peer effects modeling.

Our research adds to the growing (but currently scarce) literature focused on peer effects in the classroom. In a similar vain to the roommate peer effects literature, much of the literature on peer effects in the classroom exploits quasi-random classroom assignments. For instance, Kang (2007) exploit the quasi-random classroom assignments in South Korean middle schools, and Carman and Zhang (2012) utilize the fact that students are assigned to Chinese middle schools either randomly or through an admissions test. Both studies focused on determining the significance of peer effects on classroom grades, and find evidence of peer effects. Kang finds that strong students[2] have a positive impact on the academic performance of other students, while weak students have a negative effect on the performance of other students. Carman and Zhang find somewhat mixed results, but in general they find evidence of positive and significant peer effects on the impact of classmate quality on the academic performance of other students. Other studies use econometric techniques, such as fixed effects modeling, to correct for any selection bias. For example, Schlosser et al. (2008) and Lavy et al. (2012) control for fixed effects when studying the peer effects of underachievers and overachievers on other students in secondary schools. Both studies find evidence of peer effects where a larger proportion of lower achieving students in a high school class has a negative impact on the achievement of "regular" students in the class. Lavy et al. in particular find a positive peer effect from high achieving peers on girls. Additionally, Burke and Sass (2013) use fixed effects modeling to analyze the peer effects in Florida public school classrooms. In general, Burke and Sass find results in line with the majority of other studies, an increase in the fraction of relatively higher achieving peers makes other students better off and an increase in the fraction of relatively lower achieving peers makes other students worse off. However, there

[2] For our purposes, any qualitative measure of students abilities (high or low achiever, strong or weak student, etc.) refers to students relative acadmeic performance (usually on standardized tests). This is standard practice in all of the related literature (Carman and Zhang, 2012; Burke and Sass, 2013; Kang, 2007; Schlosser et al., 2008; Lavy et al., 2012)

are a few notable exceptions. Burke and Sass find that high achieving students benefit from an increased fraction of low or high achieving students, and in middle school mathematics, middle achievers benefit from a higher fraction of low achievers. The authors suggest that the intellectual distance between high achievers and low achievers may be the reason for their former finding. They explain that the intellectual distance may be so vast as to be a barrier to communication between the two groups, whereas high achievers can communicate effectively with the less intellectually distanced middle achievers (making high achievers worse off). Unfortunately, Burke and Sass do not provide an explanation for their later finding which resembles one of the results of this paper. With the few exceptions found by Burke and Sass (2013), the majority of the literature finds that a larger proportion of higher achievers in a class has a positive impact on the academic performance of other students, and a larger proportion of low achievers in a class has a negative impact on the academic performance of other students.

Whereas the majority of other studies focus on K-12 education, we add to the literature by analyzing the peer effects in higher education. It is possible that the peer dynamics change in the transition from secondary education to higher education, and our data set gives us a unique opportunity to analyze these peer dynamics.

3 Data

The data set comes from a selective medium sized liberal arts college in the mid-west, henceforth referred to as the institution, for which data is collected on five cohorts (2011-2015). Classroom level data is used in our regressions as it has been shown to generate stronger results than other group level measures (Burke and Sass, 2013).

In total there are 1,412 observations and 18 variables. Additionally, two different types of data are used in this study. The first type is classroom data, that consists of individual student level characteristics as well as classroom characteristics. We use data from the first class taken by first year students in order to mitigate selection bias[3]. The second type of data we use is point data, which consists of the number of points students bid on classes and their preferences for certain classes.

3.1 Classroom Data

The classroom level data on first year students in their first class at the institution is primarily used in the outcome equation.[4] There are five key variables in our outcome equation, Grade, AcadRating (Academic Rating), PctTopX, PctBotX, and PctMidY. The dependent variable, Grade is the primary outcome of interest, and is the grade received by a student after taking the course. Courses are graded on a four point scale, and there are eleven possible grades ranging from an "A" (4.0) to "F" (0.0).[5] As a measure of student ability we follow the literature and use a proxy measure developed on data prior to college enrollment, Academic Rating (Griffith and Rask, 2014; Smith and Stange, 2015). The Academic Rating is a number assigned to all students at the institution. It is a number that represents the culmination of a student's high school GPA, their test scores, the difficulty of the high school curriculum, the quality of their high school, and their writing ability. As suggested by Betts and Morell (2003) and Dooley et al. (2012) these variables are all common (and "good") indicators of college academic performance. The Academic Rating variable ranges from a low of one (representing a low ability student) to a high of sixty-five (representing a high ability student).

The Academic Rating variable is used to create our peer measure variables PctTopX, PctBotX, and PctMidY. PctTopX represents the proportion of students in the class that are in the top X percent of the sample based on Academic Rating. For instance, PctTop5 represents the proportion of students in the class that are in the top five percent of the sample based on Academic Rating. PctBotX is defined similarly for the bottom X percentage

[3] Discussed in the Empirical Methodology section.
[4] Some classroom level data is also used in the selection equation. See section 3.2 (Points Data) for details.
[5] The possible grades are A = 4.0, A- = 3.7, B+ = 3.3, B = 3.0, B- = 2.7, C+ = 2.3, C = 2.0, C- = 1.7, D+ = 1.3, D = 1.0, F = 0.0

in terms of Academic Rating. PctMidY is the proportion of students in a class that are not in the top X percent or bottom X percent of the sample in terms of Academic Rating. The remaining variables, Minority, Female, InState, Intl (International), Needy, ClassSize, URM (Underrepresented Minority), Year, Division, and Professor, are used as control variables in the regression. Table 3.1 displays the definitions of all the class level variables used.

Table 3.1: Classroom Variable Definitions

Variable	Definition
Grade	The grade received by a student after taking the course.
AcadRating	Referred to as Academic Rating, a number that represents the culmination of a student's high school GPA, their test scores, the difficulty of the high school curriculum, the quality of their high school, and their writing ability.
PctTopX	The proportion of students in the class in the top X percent of the sample based on Academic Rating
PctBotX	The proportion of students in the class in the bottom X percent of the sample based on Academic Rating
PctMidY	The proportion of students in a class that are not in the top X percent or bottom X percent of the sample in terms of Academic Rating.
Minority	A dummy variable representing whether or not the student is non-Caucasian. 1 = non-Caucasian & 0 = Caucasian
Female	A dummy variable representing whether or not the student identifies as a female. 1 = female & 0 = male
InState	A dummy variable indicating whether or not the student is an in-state student. 1 = in-state & 0 = out of state
Intl	A dummy variable indicating whether or not the student is an international student. 1 = international & 0 = not international
Needy	Whether or not a student qualified for need based financial aid. 1 = financial aid & 0 = no financial aid
ClassSize	An integer the represents the total number of students in a class.
URM	Under Represented Minority, a dummy variable indicating whether or not the student is non-Caucasian or Asian. 1 = Asian or Caucasian & 0 = other ethnicity
Year	The year the class took place.
Division	The subject area of the class, either Natural Science, Social Sciences, or Humanities.
Professor	An identifier for the professor teaching the course.

3.2 Points Data

At this institution a bidding system is used to ration classes to students. From 2011-2014 the bidding system was as follows: students are allotted twenty points and must rank eight classes in terms of their preferences, after classes are ranked students then must bid a number between zero and twenty (inclusive) points per class on their list, students with the highest number of points bid per class are allotted seats, and ties are broken randomly. If a student does not make it into any class on his or her preference list, then a class is chosen for the student at random.

In the year 2015, the bidding system was changed in an effort to allow more students to select into a class higher on their preference list. The system was changed in the following ways: students are allotted 100 points and must rank eight classes in terms of their preferences, after classes are ranked students then must bid a number between one and twenty (inclusive) points per class on their list. The remaining rules from the original system are the same. This new system effectively forces students to spread their points into multiple classes (whereas in the original system all points could be placed into one class).

The changes to the bidding system affect one of our key variables, Demand. The Demand variable is an exclusion restriction in our selection equation that represents the total number of points bid on a course divided by the number of bidders. In an attempt to correct for the changes in the bidding system, the 2015 Demand calculations are divided by five, because students receive five times the number of points compared to the original system. This corrects the mean of Demand in 2015, however an effect on standard deviation persists.

Another key variable in our selection equation is Ranking. Ranking is the dependent variable in the selection equation, and is a number between one and eight that specifies the student's preference for the course, where one is a high preference, eight is a low preference, and preferences are not repeated. The remaining variables in the first stage regression, Minority, Female, InState, Intl (International), Needy, AcadRating, URM

(Underrepresented Minority), and Subject are used as control variables. Refer to Table 3.2 for the definitions of the unique variables used in the first stage regression.

Table 3.2: Points Data Variable Definitions

Variable	Definition
Ranking	A number between one and eight specifying the student's preference for the course, where one is a high preference and eight is a low preference.
Demand	The total number of points bid on a course divided by the number of bidders. For the year 2015, this variable was divided by five to correct for the bidding system changes.
Subject	The specific subject of the course, such as mathematics, anthropology, chemistry, psychology, etc. For a full list of subjects see Appendix A.

Table 3.3 shows the number of students who selected into their first choice course, second choice course, third choice course, etc. It is important to note that the majority, 64%, of students select into their first choice course, while only 36% of students selected into a course that is not their first choice.

Table 3.3: Student Course Selections (Ranking)

Item	Number	Per cent
First Choice	903	64
Second Choice	201	14
Third Choice	136	10
Fourth Choice	62	4
Fifth Choice	38	3
Sixth Choice	36	3
Seventh Choice	16	1
Eighth Choice	20	1
Total	1,412	100

3.3 Summary Statistics

The descriptive statistics for the non-dummy variables are displayed in Table 3.4.[6] The outcome of interest, Grade, has a mean that changes slightly over time, while the standard deviation remains fairly constant. From 2011-2012 the mean Grade was 3.115, then from 2013-2015 the mean grade increased to 3.15. This suggests that there might have been some grade inflation over the years as the mean Academic Rating, a measure of student ability, remained fairly consistent from 2011-2014 (jumping by about two points in 2015). The variables PctTopX, PctBotX, and PctMidY all vary slightly from their expected values, indicating that the distribution of abilities is not uniform every year. That is, one would expect the mean of PctTop5 to always be about 0.05, however this is not the case since the Academic Rating cutoffs are not exact[7] and the distribution of abilities is not uniform. In fact, the distribution of abilities seems to be biased towards recent years, meaning the majority of high achieving students were enrolled in more recent years, as the higher mean Academic Rating, PctTop5, and PctTop10 in 2015 indicate. ClassSize jumps from a mean of 11 in 2011-2014 to 14.27 in 2015 because fewer classes were offered and more students were in the incoming class. The average number of classes offered fell from 29 in 2012-2014 to 25 in 2015.[8] As expected, the Demand variable has a lower standard deviation in 2015 compared to 2011-2014 because of the bidding system changes.

[6] Control variables, such as Year, Professor, Division, and Subject where not summarized.
[7] Exactly 5% of students do not have an Academic Rating higher than our cutoff. Instead the number is about 0.048% and this is true for all of our defined cutoffs.
[8] In 2011 the mean class size was also 25, however there were fewer students in the incoming class.

Table 3.4: Summary Statistics

	2011 Data		2012 Data		2013 Data		2014 Data		2015 Data		Overall	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Grade	3.12	0.96	3.11	0.87	3.17	0.78	3.15	0.81	3.16	0.87	3.14	0.86
AcadRa~g	50.72	6.09	50.89	5.9	49.98	6.25	50.95	6.44	52.80	6.19	51.12	6.26
PctTopQ	0.30	0.13	0.28	0.17	0.22	0.17	0.27	0.15	0.41	0.14	0.30	0.17
PctTop5	0.08	0.10	0.06	0.10	0.06	0.07	0.04	0.07	0.14	0.10	0.08	0.10
PctTop10	0.16	0.12	0.14	0.13	0.14	0.11	0.16	0.13	0.26	0.13	0.18	0.13
PctBotQ	0.23	0.11	0.19	0.11	0.28	0.17	0.18	0.10	0.13	0.10	0.20	0.13
PctBot5	0.04	0.06	0.05	0.08	0.06	0.12	0.06	0.08	0.04	0.06	0.05	0.08
PctBot10	0.08	0.08	0.07	0.08	0.10	0.14	0.08	0.08	0.06	0.08	0.08	0.09
PctMid50	0.47	0.16	0.54	0.16	0.50	0.17	0.54	0.14	0.46	0.14	0.50	0.16
PctMid90	0.87	0.11	0.89	0.12	0.87	0.14	0.90	0.12	0.82	0.10	0.87	0.12
PctMid80	0.77	0.14	0.78	0.14	0.76	0.16	0.76	0.16	0.68	0.13	0.75	0.15
Minority	0.21	0.41	0.37	0.48	0.46	0.50	0.35	0.48	0.39	0.49	0.36	0.48
Female	0.47	0.50	0.55	0.50	0.53	0.50	0.51	0.50	0.54	0.50	0.52	0.50
InState	0.15	0.35	0.13	0.33	0.15	0.36	0.12	0.33	0.15	0.36	0.14	0.35
Intl	0.02	0.15	0.01	0.08	0.06	0.23	0.05	0.22	0.07	0.25	0.04	0.20
Needy	0.39	0.49	0.36	0.48	0.36	0.48	0.32	0.47	0.43	0.50	0.37	0.48
ClassS~e	10.85	3.09	11.91	3.05	11.11	2.48	11.39	2.62	14.27	2.47	12.00	3.01
URM	0.11	0.31	0.29	0.45	0.36	0.48	0.27	0.44	0.30	0.46	0.27	0.45
Ranking	1.77	1.54	1.70	1.44	1.79	1.41	1.91	1.79	2.14	1.41	1.87	1.53
Demand	2.81	0.83	2.62	0.75	2.28	0.86	2.37	0.86	2.44	0.22	2.49	0.76
N	226		292		290		285		319		1,412	

Our data gives us a unique opportunity to study peer effects in higher education. First, our data comes from an institution with relatively small class sizes which Diette and Raghav (2015) and Kokkelenberg et al. (2008) show have a positive impact on average student achievement (measured by grades and test scores). The primary reasoning for the inverse relationship between achievement and class size is that students have more quality time to interact with teachers and peers as class size decreases, suggesting that we may expect to find magnified peer effects at this institution. Additionally, the data on student preferences (Ranking) and class demands (Demand) enable us to use a two stage selection model, a model generally not used in other classroom peer effects literature[9], to combat selection bias and analyze the data.

4 Empirical Methodology

Our foundational peer effects model is as follows:

$$G_i = \beta_0 + \beta_1 Ability_i + \beta_2 Ability_i^{CM1} + \beta_3 Ability_i^{CM2} + \overrightarrow{\beta}\,\overrightarrow{z} + \varepsilon_i \tag{1}$$

This is an OLS model, where G_i is the grade received by student i in their first course at the institution, $Ability_i$ is a proxy for the student's academic ability (Academic Rating)[10], $Ability_i^{CM1}$ and $Ability_i^{CM2}$ are classmate ability measures, $\overrightarrow{z}$ is a vector of control variables[11], and ε_i is the error term. $Ability_i^{CM1}$ and $Ability_i^{CM2}$ are our peer measures and are defined as one of the following, the proportion of high achieving, middle achieving, or low achieving students in a class[12](as defined by cutoffs in academic rating), and $Ability_i^{CM1}$ is not the same measure as $Ability_i^{CM2}$. Thus β_2 and β_3 are the primary coefficients of interest, as they estimate the impact of the classmates ability variables

[9] See section 2 (Literature Review) for more details.
[10] For more information see Table 3.1.
[11] Control variables included Minority, Female, InState, International, Needy, Class Size, URM, Year, Division, and Professor. See Table 3.1 and Table 3.2 for definitions.
[12] Unfortunately due to data limitations only the proportion of high achievers and low achievers in a class is used for our peer measures, for more details see section 5 (Results).

(our peer measures) on a student's grade. The OLS regression is run on the subset of students that are not incorporated in the peer measures. For example, if $Ability_i^{CM1}$ and $Ability_i^{CM2}$ are defined as the proportion of high achievers in a class and the proportion of low achievers in a class respectively, then the model would be run on the middle achievers.

However, because our sample is nonrandom[13] the calculated coefficients of model (1) are at risk of being bias (Heckman, 1979).[14] This bias, known as selection bias, is a serious threat to peer effects models of this type because self selection often exists and leads to a nonrandom sample (Carman and Zhang, 2012; Burke and Sass, 2013; Ding and Lehrer, 2007). Therefore, we must update our model to correct for any selection bias that may exist.

4.1 Two Stage Selection Model

As discussed in section 2 (Literature Review), most studies of this kind exploit quasi-random student class assignments or utilize fixed effects models to control for selection bias (Kang, 2007; Carman and Zhang, 2012; Schlosser et al., 2008; Lavy et al., 2012). However, due to our unique access to student preference data, we use a two stage selection model to correct for selection bias, similar to the one described in Heckman (1979). Our two stage selection model uses an ordered probit model in the first stage (the selection equation) and an OLS model in the second stage (the outcome equation). From the first stage ordered probit model we take the calculated inverse mills ratios and use them as a control variable in the second stage outcome equation. The inverse mills ratios are calculated estimates, that when used in the outcome equation, help to control for selection bias (Heckman, 1979).[15]

---

[13] Unfortunately students are not assigned to classes and institutions randomly, instead the institution selects specific students, students next select the institution, and students then select into classes.
[14] For more details see section 4.2 (Controlling Selection Bias)
[15] For more information see Greene (2002).

For the selection equation we use an ordered probit model, defined as follows:

$$R_i^* = \alpha_1 D_i + \overrightarrow{\alpha}\,\overrightarrow{\omega} + \varepsilon_i \tag{2}$$

$$R_i = \begin{cases} 1 \;\; if \;\; -\infty < R_i^* \leq \mu_1 \\ \vdots \\ j \;\; if \;\; \mu_j < R_i^* < \infty \end{cases} \tag{3}$$

Where D_i is the demand for the student's selected class, $\overrightarrow{\omega}$ is a vector of control variables[16], and ε_i is the error term. In (3) we see that the unobserved selection variable R_i^* corresponds to the observed R_i through μ, a vector of unknown cutoffs. The variable j represents the number of selection categories there are, where in any of our regressions j is at least two and at most four. For example, if j is equal to two, then R_i equal to one represents those students who selected into their first choice class and R_i equal to two represents those students who did not select into their first choice class (instead selecting into their second choice, third choice, etc.). The ordered probit model ultimately estimates the probability that R_i is equal to j using R_i^*, that is,

$$Pr(R_i = j) \;=\; Pr(\mu_{j-1} < R_i^* \leq \mu_j) \tag{4}$$

Once we run the selection equation, we use the inverse mills ratios in the outcome equation.

We use an OLS model very similar to our foundational model as our outcome equation. The only difference is that the inverse mills ratios are added as a control variable to correct for selection bias. The model is as follows:

$$G_i = \beta_0 + \beta_1 Ability_i + \beta_2 Ability_i^{CM1} + \beta_3 Ability_i^{CM2} + \beta_4 \lambda_i + \overrightarrow{\beta}\,\overrightarrow{z} + \varepsilon_i \tag{5}$$

Where the variables are defined in the same manner as (1), and λ_i is the inverse mills

[16] Specifically the variables include Minority, Female, InState, Intl, Needy, AcadRating, URM (Underrepresented Minority), and Subject. See Table 3.1 and Table 3.2 for definitions.

ratios taken from the first stage regression. Just as in (1) the OLS regression is run on the subset of students that are not part of the peer measures. However, the calculation of different inverse mills ratios for each of the possible selection categories makes it necessary for a separate regression to be run on each of the students who selected into a particular selection category.[17] As an example, suppose $Ability_i^{CM1}$ and $Ability_i^{CM2}$ are defined as the proportion of high achievers in a class and the proportion of low achievers in a class respectively and the number of selection categories (j) is equal to two. The outcome equation (5) will first be regressed on the middle achievers who selected into their first choice class, then on the middle achievers who did not select into their first choice class, producing two sets of regression outputs.

4.2 Controlling Selection Bias

We use two methods in an attempt to control for selection bias, a serious problem in many peer effect models (Carman and Zhang, 2012; Burke and Sass, 2013; Ding and Lehrer, 2007). The first method, as mentioned in section 3 (Data), is to limit our data to first year students in their first course at the institution. The incoming classes at the intitution come from a diverse background (both geographically and socioeconomically)[18] so it is highly unlikely that students are selecting their first course based on any prior relationship to classmates. This effectively eliminates one possible area of selection bias, self-selecting into courses based on prior bonds/relationships.

The second method we use to correct for selection bias is the two stage selection model. Specifically, our model is controlling for any selection bias that results from students selecting into classes high on their preference list or students selecting into classees low on their preference list.[19] The idea is that there are unobservable factors that correlate with both student class preferences and the grade received in the class, which would then bias

[17] This is why there are Grade1, Grade2, etc. categories in the regressions seen in section 5 (Results).
[18] Compared to other institutions of this type.
[19] The cuttoffs that define a “high” and “low” preference course are determined by the number of categories (j) in our selection model.

our coefficients of interest (β_2 and β_3). For instance, it may be that popular courses draw a higher proportion of high achieving students, while the majority of middle achieving students are drawn into those same courses because they enjoy the subject. In this case a higher proportion of high achievers is not causal to the increased grades of middle achievers. Instead, an unobservable variable (passion for the subject) drives middle achievers to self select into courses where they will achieve higher grades and there happen to be a higher proportion of high achievers. In such a case, the self-selection of middle achieving students would bias our coefficients of interest and our results would be inaccurate. By using the two stage regression model, we are attempting to correct for this type of selection bias.

5 Results

The results suggest that the proportion of high achievers in a class has a significant, but small, negative impact on the grades of middle achievers, while the proportion of low achievers in a class has a significant, but small, positive impact on the grades of middle achievers[20]. We find that the estimated coefficients are too small to have an economically significant impact on final student grades, however if these trends continue in the long run they may impact cumulative GPA. These results are contradictory to those found in the majority of the literature and we discuss some possible explanations for this below (Kang, 2007; Carman and Zhang, 2012; Burke and Sass, 2013; Schlosser et al., 2008; Lavy et al., 2012). We begin the section by evaluating our results in detail and describing the patterns found. Then, we conclude by discussing the implications of our results and provide some suggestions for further research.

It is important to note that the main results of this paper focus on the impact of high achievers and low achievers on the grades of middle achievers. Unfortunately, data limitations made it difficult to measure the impact of classmate ability on high achievers

[20] See Table 5.1, Table 5.2, and Table 5.3

or low achievers.[21] There are not enough, observations at this time, within either group (high achievers and low achievers) to run the two stage regression model. However, the regressions run on the middle achievers still make an important contribution as the majority of students are middle achievers, so our findings impact most of the population.

Table 5.1, Table 5.2, and Table 5.3 show the relevant results[22] of the two stage selection model (outlined in section 4) run on the middle achievers, where the selection equation (2) is a two category ordered probit, three category ordered probit, and four category ordered probit respectively. Each table contains the relevant results from three regressions where the major difference between each regression within a table is the measure of the "peer" used. For the regressions found in Table 5.1, 5.2, and 5.3, we use the proportion of high achievers and low achievers in a class as a peer measure and regress on middle achieving students. After each regression, the measures used for a high and low achieving peer become stricter (and therefore the measure for the middle achievers becomes less strict). For instance, regression 1 (labeled Top/Bottom 25%) in Table 5.1 focuses on the impact of the proportion of high achievers, defined as the twenty-five percent with the highest academic rating in the sample, and the proportion of low achievers, the twenty-five percent with the lowest academic rating in the sample, on the grades of the middle achievers (those that are not high or low achievers). In each regression, GradeX refers to the regression run on the middle achievers that selected into their X choice class based on the first stage ordered probit.[23]

[21] Specifically we could not measure the impact of the proportion of low achievers and middle achievers on high achievers and similarly for low achievers.

[22] The results of the regression for all the variables can be found in Appendix B.

[23] For more details see section 4.1 (Two Stage Selection Model)

Table 5.1: Two Selection Categories

	Top/Bott~25% (Std. Err.)	Top/Bott~10% (Std. Err.)	Top/Bott~5% (Std. Err.)
Grade1			
PctBotQ	-0.150		
	(0.53)		
PctTopQ	-0.787**		
	(0.41)		
PctBot10		0.612	
		(0.54)	
PctTop10		-0.741***	
		(0.33)	
PctBot5			1.001**
			(0.60)
PctTop5			-0.702**
			(0.41)
Grade2			
PctBotQ	1.720**		
	(1.00)		
PctTopQ	0.660		
	(0.70)		
PctBot10		0.370	
		(0.78)	
PctTop10		0.827**	
		(0.49)	
PctBot5			-0.061
			(0.70)
PctTop5			0.198
			0.198

* p<.2, ** p<.1, *** p<.05

Note: Regressions run using 3 different peer measures and a two choice ordered probit for the first stage.

Table 5.2: Three Selection Categories

	Top/Bott~25% (Std. Err.)	Top/Bott~10% (Std. Err.)	Top/Bott~5% (Std. Err.)
Grade1			
PctBotQ	-0.314		
	(0.52)		
PctTopQ	-0.840***		
	(0.39)		
PctBot10		0.602	
		(0.54)	
PctTop10		-0.705***	
		(0.33)	
PctBot5			1.008**
			(0.60)
PctTop5			-0.667*
			(0.41)
Grade2			
PctBotQ	1.248		
	(1.41)		
PctTopQ	1.614**		
	(0.92)		
PctBot10		-0.138	
		(1.02)	
PctTop10		0.439	
		(0.70)	
PctBot5			-0.819
			(0.88)
PctTop5			-0.585
			(1.00)
Grade3			
PctBotQ	1.215		
	(1.54)		
PctTopQ	-0.961		
	(1.16)		
PctBot10		0.002	
		(1.18)	
PctTop10		0.541	
		(0.63)	
PctBot5			-0.170
			(1.13)
PctTop5			-0.031
			(0.83)

* p<.2, ** p<.1, *** p<.05

Note: Regressions run using 3 different peer measures and a three choice ordered probit for the first stage.

Table 5.3: Four Selection Categories

	Grade1		Grade2		Grade3		Grade4	
	PctBot5	**PctTop5**	**PctBot5**	**PctTop5**	**PctBot5**	**PctTop5**	**PctBot5**	**PctTop5**
Top/Bott~5%	0.998**	-0.669*	-0.791	-0.530	2.443*	-0.215	-0.417	-2.847***
(Std. Err.)	(0.61)	(0.41)	(0.88)	(1.01)	(1.64)	(1.52)	(1.72)	(1.18)

* $p<.2$, ** $p<.1$, *** $p<.05$

Note: Regressions run using 3 different peer measures and a three choice ordered probit for the first stage. Additionally, Top/Bottom 25% and Top/Bottom 10% regressions could not be run because there is too little variance in the data at those levels (most likely because there are too few observations).

When analyzing the results in Table 5.1, 5.2, and 5.3 there are a few important patterns to note. First, the majority of the statistically significant results are found in students who selected into their first choice (Grade1), and students who selected into their second choice or below (Grade2, Grade3, etc.) show inconsistent results. Next, all the significant results in Grade1 are consistent, in that the proportion of high achievers has a negative impact on the grades of middle achievers and the proportion of low achievers has a positive impact on the grades of middle achievers. Furthermore, the majority of significant results appear as the peer measure becomes stricter, that is more significant results appear as the peer measure changes from the top/bottom twenty-five percent, to ten percent, and finally to five percent.

The Grade1 results in Table 5.1, 5.2, and 5.3 show a consistent pattern. The proportion of high achievers in a classroom has a negative impact on the grade of middle achievers, regardless of the strictness of the peer measure. As the peer measure becomes stricter (specifically at the five percent threshold), we see that the proportion of low achievers in a classroom has a positive impact on the grades of middle achievers. There is some precedence for using the five percent threshold and finding peer effects. In Lavy et al. (2012) researchers also found strong peer effects using the top and bottom five percent as their peer measures. Lavy et al. argue that their use of the top and bottom five percent is not arbitrary by showing that it is precisely those students around the five percent threshold that have a strong peer impact on fellow students. That is, they showed that students in the middle ninety percent of the ability distribution do not show strong peer effects of any sort, while the top and bottom five percent are the students who are the most influential.

Our results run counter to those found in the current literature[24], granted that the current literature is rather scarce especially in the area of higher education (Kang, 2007; Carman and Zhang, 2012; Burke and Sass, 2013; Schlosser et al., 2008; Lavy et al., 2012). Most of the current literature finds that low achievers hurt the grades of middle achieving students

[24] Burke and Sass (2013) do find one result similar to ours, but they do not provide any explaination for their finding.

and high achievers help the grades of middle achieving students. Similar results have also been discovered in roommate peer effects literature (Griffith and Rask, 2014; Zimmerman, 2003; Sacerdote, 2000). The reasoning being that low achieving students may disrupt learning (possibly by instilling bad study habits in the middle achieving students) and the high achieving students may facilitate learning (possibly by asking more relevant questions or helping to tutor middle achieving students). This reasoning is very compelling and is exactly what many people would expect to happen, which makes our results of the contrary all the more puzzling.

We have developed two possible explanations for these results. One possibility is that since most of the previous literature does not focus on courses in higher education, peer dynamics have changed in the transition from K-12 education to undergraduate education. Most of the literature focuses on K-12, therefore, it is reasonable to suppose that peer effects at institutions of higher education may differ from those seen in K-12 education. It is possible that high achieving students in college have a negative peer effect on grades, perhaps through demoralizing other students, and low achieving students have a positive peer effect on grades, perhaps through increasing the number of group study sessions. Another possibility is that grade curving is occurring in many of the classes, and is overpowering any actual peer effects. That is to say, a student's grade in a class may be determined by her relative performance to her classmates instead of by her absolute performance. Unfortunately, due to the time span of the data and faculty changes, it is not feasible to uncover which classes were truly graded on a curve (in which case these findings would not be surprising) and which classes were not. Furthermore, if grade curving is the underlying cause of the results we are finding, then it is still possible that peer effects that are in line with the literature exist, but the curving effects are simply overwhelming any peer effects that are occurring.

Next, we'll explore the possible reasons behind the inconsistent results in the regressions run on students who selected into a class that was not their first choice. One reason, may be that for students who do not select into their first choice class, the peer effect dynamics change and aren't as strong. Perhaps the students who did not select into their first choice class have less motivation (or something of the sort) to perform in their non-first choice class, and this leads to less peer interactions (or interactions of a different kind) that ultimately lead to smaller/immeasurable peer effects. Another explanation, one that we believe is more likely, is that there are simply fewer students who selected into a course that is not their first choice and the results are therefore less reliable. In the best cases the number of students who did not select into their first choice is about one half of the number of students who did select into their first choice.[25] That is, for a two option ordered probit selection model (used in Table 5.1), the number of students who are regressed in the Grade1 calculations (those who selected into their first choice class) is twice that of the number of students are regressed in the Grade2 calculations (those who did not select into their first choice class).[26] Additionally, as the number of options in the selection model increases (as is the case in Table 5.2 and 5.3) the number of students who did not select into their first choice class is further divided into those who selected into their second choice class, third choice class, etc. and regressed on. This essentially means that the regressions run on the students who did not select into their first choice class (Grade2, Grade3, Grade4) have far fewer observations than the regressions run on students who did select into their first choice class (Grade1). In this case, fewer observations effectively means a smaller sample size, and studies that use smaller sample sizes find that the sample size may lead to insignificant or inaccurate results (González-Val et al., 2013) and the results of these studies are often not generalizable (Oladipupo and Ibadin, 2013). This is a troubling problem as it suggests that the results from the Grade2, Grade3, and Grade4 regressions are more unreliable than the results of the Grade1 regressions. However, the only way to resolve this problem is to collect several more years of data, which is not feasible at this time. Therefore, we will focus on the Grade1 regression results for the

[25] See Table 3.3 in section 3.2 (Points Data) for more details
[26] For more details see section 4 (Empirical Methodology).

remainder of the section, and since the majority of students selected into their first choice class, they are arguably the most relevant results regardless.

Additionally, our results are somewhat imbalanced. That is to say, high achieving students clearly have a negative impact on middle achieving students throughout the majority of our regressions, but low achieving students only have a statistically significant positive impact on middle achieving students when the peer measure becomes stricter. One explanation may be that it is pure statistical chance that high achievers had a significant negative affect more often than low achievers had a significant positive affect. Our sample size of 903 students who selected into their first choice class is not as large as some found in the literature, which reach into the many thousands (Kang, 2007; Lavy et al., 2012). Another possibility is that top students are very effective at having a negative impact on other student's grades, and this result is simply robust to changes in the peer measure. Finally, if grade curving is the underlying cause of our results, it may be that high achieving students dull out a middle achieving student's performance more than a low achieving student helps a middle achieving student shine. Since professors may curve grades in a subjective fashion, they may be affected by this psychological phenomenon. Currently we do not have an explanation for this phenomena, although there is evidence of similar kinds of psychological imbalance in other areas. For example take the well documented theory of loss aversion. As explained by Tversky and Kahneman (1991), loss aversion implies that people experience greater impact from a loss when compared to a gain equal in magnitude, and this phenomena has been found in many different areas of human behavior (Shalev, 2002; Goette et al., 2004). It is possible that the same psychological mechanisms driving the imbalance in losses and gains (loss aversion) are also driving the imbalance in the effect of the proportion of high achievers and low achievers on middle achievers grades.

There are two concerns with the methodology used in this paper that need to be addressed. One concern is selection bias and the other is the change in 2015 Demand[27] which may affect the model and overall results. In order to correct for selection bias we use the two stage selection model outlined in section 4.1 (Two Stage Selection Model), and after comparing our results to the foundational OLS model outlined in (1), the overall results are the same.[28] This suggests that selection bias is most likely not affecting our sample to any significant degree, but selection bias may still be a concern with any future studies that utilize a different sample. To address our second concern, we correct Demand in 2015[29] to eliminate some of the effects of the changes in the bidding system, however some effects persisted[30] and may have affected our results. In order to verify the robustness of our results we reran the regressions on the same data from the years 2011-2014, with the original bidding system, and found that the same overall results persist.[31] Therefore neither of our initial concerns had a significant affect on the overall results of the paper.

5.1 Discussion

The above results indicate the presence of an interesting trend at this institution, but are there any significant implications from these trends? In order to shed some light on the implications of our results, we must analyze the magnitudes of the estimated coefficients. The average class size in our data set is twelve students (which has remained fairly constant throughout the time period of our data set), and if we define high achievers as the top five percent of students then there is one student per class that is a high achiever, assuming a uniform distribution of high achievers. At the five percent threshold, our results indicate that for every ten percent increase in the proportion of high achievers in the classroom the GPA of a middle achiever falls by roughly 0.07 points. In order for the proportion of high

[27] See section 3.2 (Points Data) for more details.
[28] See Appendix C for the regular OLS results.
[29] See section 3.2 (Points Data) for more details on the correction methods we used.
[30] In particular the standard deviation of Demand in 2015 was still affected. See section 3 (Data) for more details.
[31] TODO: See appendix X for the full results.

achievers in an average class to be high enough to have a significant[32] influence on the grades of middle achieving students, forty percent of the average class (five students) must be high achievers.[33] Using the five percent threshold, there are two classes (about 1.5%) in our entire data set which have a proportion of high achievers of at least forty percent. It is clear that having a class consist of at least forty percent high achievers is an unlikely scenario, and should not be a practical concern for most students. Similar calculations can be done to show that having a high enough number of low achieving students in a classroom to significantly affect grades is also quite uncommon. However, two considerations must be taken into account. First, if these results continued in further classes, then the cumulative effect on the GPA may be significant. Depending on a student's degree path, they may be more likely to select into classes with a higher proportion of low achieving students or classes with a higher proportion of high achieving students. Additionaly, higher level classes at the institution may number under ten students (depending on the popularity of the degree), therefore if these trends continue it is possible for the cumulative GPA (as well as class grades) of a student to be affected. This may be a serious problem in unpopular degrees with abnormally small class sizes. The second consideration is that the coefficients are estimates, and it may be that the real coefficients are much larger or smaller than those estimated. If the real coefficients happen to be larger, then the results may indeed be a practical concern to students and the administration.

The calculations above show that even though these trends exist, in the short run, the impact on final grades in the first course is most likely not a practical concern to most students. However, more research need to be conducted to determine the long run implications of these results.

There are two types of policies that the majority of the literature discusses, "streaming" and "mixing" (Ding and Lehrer, 2007; Kang, 2007; Carman and Zhang, 2012). Streaming refers to the policy of grouping students of similar ability levels, while mixing generally

[32] Defined as at least changing the sign of a student's grade which takes about 0.3 grade points.

[33] $0.4 * 0.7 = 0.28$

refers to the policy of combining students of different ability levels (Ding and Lehrer, 2007). However, in this case it is difficult to suggest any policy implications without restricting student course choices and without knowing how high achievers and low achievers are affected by classmates of different ability levels.

As Burke and Sass (2013), Carman and Zhang (2012), and Ding and Lehrer (2007) find, an optimal policy prescription may not exist. Suppose that the grades of high achievers benefit from middle and low achievers, while the grades of low achievers are harmed by middle and high achievers,[34] then a trade off will need to be made if a policy is to be implemented. In this case, there is no grouping which would lead to a positive impact on all student grades. For instance, if middle achievers are prioritized, since there are more of them, then they should be placed into classes with as many low achievers as possible. However, this leads to two problems. First, the grades of low achievers would be harmed in this case. Additionally only a select group of middle achievers could benefit from this policy, since there are far fewer low achievers. Also, if too many middle achievers are placed into a class, the proportion of low achievers would be diluted to a point where the affects are insignificant. If, on the other hand, the grades of low achievers are positively impacted by the proportion of middle achievers in the class, then by grouping low achievers and middle achievers the grades of both may be maximized. Again, only a select group of middle achievers would benefit, but this would still maximize the average grade received by all groups overall. However, to effectively create this optimal grouping, the institution would need to decide to restrict student course choices based on their ability levels, which may not be a popular decision among current or prospective students.[35]

This study is somewhat limited by the sample size. With more observations we may be able to determine how high achievers and low achievers are affected by each other and middle achievers. Additionally, it is an open question as to whether or not these results

[34] Which is what Appendix C suggests to be the case.

[35] Institutions could also incentivise students of certain ability levels to take certain courses, however this may be considered unfair to those students who are not incentivised.

may be generalized to other types of institutions. This study was conducted on data from a small private liberal arts college and the results may differ at different types of institutions with potentially different peer interactions (such as at large universities). Also, due to the time span of the data, we are unable to determine whether peer effects, grade curving, or both are responsible for our results.

Fortunately, these limitations may lead to promising areas of future research. In a few years there may be enough observations to rerun the regressions on high and low achievers and discover how they are affected by students of other ability levels in the class. With this full picture of classroom peer effects one might find unexpected peer effects (such as if low achievers benefited from high achievers and high achievers benefited from low achievers) that are currently unexplained and have clearer policy implications. Another avenue for further research, would be to conduct similar regressions at different types of institutions. It would be interesting to know if the same peer effects are found at community colleges or large universities. Both community colleges and large universities are a fundamentally different type of institution compared to a small private liberal arts school, and it is possible that different types (or the same types) of peer effects exist. Additionally, different liberal arts colleges may have different grading policies (i.e. a specific policy on grade curving), and by comparing the results from another (similar) institution to this one it may help determine if grade curving is an underlying cause of our results. Finally, it may be interesting to analyze classroom peer effects based on gender composition. Oosterbeek and Van Ewijk (2014) study gender peer effects in a university and find that "males, but not females, perform poorer in courses with a high math component if the share of females in their work group increases." It is possible that the majority of our results are due to gender peer effects (instead of the ability peer effects that we attribute them to). That is, since high achievers are generally female (71% are female at the 5% threshold) the strongest driver of our overall results may be from the impact of a high achieving female on the grade of a middle achieving male.

6 Conclusion

Several studies have found that the quality of ones peers at an institution of higher learning has a positive impact on long term outcomes, such as graduation rates and cumulative GPA (Smith and Stange, 2015; Luppino and Sander, 2015; Ost, 2010). We seek to answer the question of whether or not peer quality matters in the short term, particularly at the classroom level, and find small but significant effects from the proportion of high achievers and low achievers in a class on the grades of middle achievers. The results indicate that the proportion of high achievers in a class has a negative impact on the grades of middle achievers, while the proportion of low achievers in a class has a positive impact on the grades of middle achievers. These results are novel and run counter to those found in the majority of the current literature (Kang, 2007; Carman and Zhang, 2012; Burke and Sass, 2013; Schlosser et al., 2008; Lavy et al., 2012), which may be due to changing peer dynamics from K-12 education to higher education, grade curving occurring in classes at the institution, or a combination of both.

Due to data limitations we are unable to uncover how classmate ability impacts the grades of high and low achievers, which makes it challenging to prescribe any practical policy recommendations. Additionally, the type of institution studied was relatively unique, and it is not clear how generalizable these results are. Fortunately, these study limitations create promising avenues for future research. Particularly promising areas of future research include comparing the classroom level peer effects at different types of institutions (such as community colleges or large universities), and, following in the footsteps of Oosterbeek and Van Ewijk (2014), looking at classroom peer effects from a gender peer effects perspective.

References

Betts, J. and Morell, D. (2003), 'The determinants of undergraduate grade point average: The relative importance of family background, high school resources, and peer group effects', *International Library of Critical Writings in Economics* **165**, 310–335.

Black, D. A. and Smith, J. A. (2006), 'Estimating the returns to college quality with multiple proxies for quality', *Journal of labor Economics* **24**(3), 701–728.

Burke, M. A. and Sass, T. R. (2013), 'Classroom peer effects and student achievement', *Journal of Labor Economics* **31**(1), 51–82.

Carman, K. G. and Zhang, L. (2012), 'Classroom peer effects and academic achievement: Evidence from a chinese middle school', *China Economic Review* **23**(2), 223–237.

Diette, T. M. and Raghav, M. (2015), 'Class size matters: Heterogeneous effects of larger classes on college student learning', *Eastern Economic Journal* **41**(2), 273–283.

Ding, W. and Lehrer, S. F. (2007), 'Do peers affect student achievement in china's secondary schools?', *The Review of Economics and Statistics* **89**(2), 300–312.

Dooley, M. D., Payne, A. A. and Robb, A. L. (2012), 'Persistence and academic success in university', *Canadian Public Policy* **38**(3), 315–339.

Foster, G. (2006), 'It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism', *Journal of public Economics* **90**(8), 1455–1475.

Goette, L., Huffman, D. and Fehr, E. (2004), 'Loss aversion and labor supply', *Journal of the European Economic Association* **2**(2-3), 216–228.

González-Val, R., Lanaspa, L. and Sanz-Gracia, F. (2013), 'Gibrat's law for cities, growth regressions and sample size', *Economics Letters* **118**(2), 367–369.

Greene, W. H. (2002), 'Limdep, version 8.0. econometric modeling guide, vol 1. plainview, ny: Econometric software', *Inc., pp. E14-9-E14-11* .

Griffith, A. L. and Rask, K. N. (2014), 'Peer effects in higher education: A look at heterogeneous impacts', *Economics of Education Review* **39**, 65–77.

Heckman, J. (1979), 'Sample selection bias as a specification error.', *Econometrica* .

Kang, C. (2007), 'Classroom peer effects and academic achievement: Quasi-randomization evidence from south korea', *Journal of Urban Economics* **61**(3), 458–495.

Kokkelenberg, E. C., Dillon, M. and Christy, S. M. (2008), 'The effects of class size on student grades at a public university', *Economics of Education Review* **27**(2), 221–233.

Lavy, V., Silva, O. and Weinhardt, F. (2012), 'The good, the bad, and the average: evidence on ability peer effects in schools', *Journal of Labor Economics* **30**(2), 367–414.

Luppino, M. and Sander, R. (2015), 'College major peer effects and attrition from the sciences', *IZA Journal of Labor Economics* **4**(1), 1–23.

McEwan, P. J. and Soderberg, K. A. (2006), 'Roommate effects on grades: Evidence from first-year housing assignments', *Research in Higher Education* **47**(3), 347–370.

Oladipupo, A. O. and Ibadin, P. O. (2013), 'Does working capital management matter in dividend policy decision? empirical evidence from nigeria', *International Journal of Financial Research* **4**(4), 140.

Oosterbeek, H. and Van Ewijk, R. (2014), 'Gender peer effects in university: Evidence from a randomized experiment', *Economics of Education Review* **38**, 51–63.

Ost, B. (2010), 'The role of peers and grades in determining major persistence in the sciences', *Economics of Education Review* **29**(6), 923–934.

Sacerdote, B. (2000), Peer effects with random assignment: Results for dartmouth roommates, Technical report, National bureau of economic research.

Sarmiento Espinel, J. A., Silva Arias, A. C. and Van Gameren, E. (2015), 'Quality differences of higher education and its determinants in a less-developed country', *Journal of Higher Education Policy and Management* **37**(2), 204–221.

Schlosser, A., Lavy, V. and Paserman, M. D. (2008), *Inside the Black Box of Ability Peer Effects: Evidence from Variation in Low Achievers in the Classroom*, Pinhas Sapir Center for Development.

Shalev, J. (2002), 'Loss aversion and bargaining', *Theory and decision* **52**(3), 201–232.

Smith, J. and Stange, K. (2015), A new measure of college quality to study the effects of college sector and peers on degree attainment, Technical report, National Bureau of Economic Research.

Tversky, A. and Kahneman, D. (1991), 'Loss aversion in riskless choice: A reference-dependent model', *The quarterly journal of economics* pp. 1039–1061.

Zimmerman, D. J. (2003), 'Peer effects in academic outcomes: Evidence from a natural experiment', *Review of Economics and statistics* **85**(1), 9–23.

Appendices

A Subjects

There were 31 subjects used as controls in the selection equation, the full list includes:

1. Art History (AH)
2. Anthropology (AN)
3. Biology (BY)
4. Chemistry (CH)
5. Classics (CL)
6. Chinese (CN)
7. Comparative Literature (CO)
8. Dance Theory (DA)
9. Dance Studio (DR)
10. Education (ED)
11. English (EN)
12. Environmental Science (ES)
13. Film Studies (FS)
14. German (GR)
15. Geology (GY)
16. History (HS)
17. Psychoanalysis (HY)
18. Japanese (JA)
19. Mathematics (MA)
20. Molecular Biology (MB)
21. Music (MU)
22. NS[36]
23. Physics (PC)
24. Philosophy (PH)
25. Political Science (PS)
26. Psychology (PY)
27. Religion (RE)
28. Russian (RS)
29. Sociology (SO)
30. Spanish (SP)
31. Theatre (TH)

B Full Regressions

All of the coefficients and standard errors from the regressions run in Table 5.1, Table 5.2, and Table 5.3 are shown below (with the exception of the professor dummy variables).

Note: For the four category two stage selection model (Table B.10 and Table B.11) only the Top/Bot 5% regression could be run. Top/Bot 25% and Top/Bot 10% regressions could not be run because there is too little variance in the data at those levels (most likely because there are too few observations).

[36] One course in 2011 was regestered with this label, but records are lossed as to what it stands for.

Table B.1: 2 Selection Categories, 2 Stage Regression on Middle Achievers using Top/Bot 25%

Variable	Coefficient	(Std. Err.)	Variable	Coefficient	(Std. Err.)
Equation 1 : Selection 2 Categories			Equation 3 : Grade1		
Minority	-0.351**	(0.197)	_lambda	-0.298	(0.289)
Female	-0.008	(0.113)	Minority	0.062	(0.117)
InState	0.053	(0.177)	Female	0.237***	(0.069)
Intl	0.226	(0.285)	InState	-0.049	(0.106)
Needy	-0.014	(0.150)	Intl	0.006	(0.188)
AcadRating	0.010	(0.025)	Needy	-0.009	(0.091)
URM	0.068	(0.209)	AcadRating	0.014	(0.015)
Demandc	-0.492***	(0.089)	PctBotQ	-0.150	(0.529)
SUB_AH	-0.599	(0.471)	PctTopQ	-0.787**	(0.405)
SUB_AN	-1.207***	(0.448)	ClassSize	-0.009	(0.021)
SUB_BY	-0.210	(0.440)	URM	-0.127	(0.116)
SUB_CH	-1.535***	(0.534)	Year1	0.000	(0.000)
SUB_CL	-0.115	(0.473)	Year2	0.210*	(0.144)
SUB_CN	-0.214	(0.498)	Year3	0.185	(0.166)
SUB_CO	-0.277	(0.477)	Year4	0.136	(0.148)
SUB_DA	-0.399	(0.721)	Year5	0.227	(0.225)
SUB_DR	0.585	(0.971)	Sub_H	0.616*	(0.403)
SUB_ED	0.055	(0.569)	Div_N	0.121	(0.417)
SUB_EN	-0.838***	(0.391)	Div_S	0.390	(0.418)
SUB_ES	-0.219	(0.766)	Intercept	2.122***	(1.021)
SUB_FS	-1.753***	(0.600)	Equation 4 : Grade2		
SUB_GR	-0.039	(0.499)	_lambda	-0.112	(0.404)
SUB_GY	-0.718*	(0.485)	Minority	-0.034	(0.197)
SUB_HS	0.032	(0.471)	Female	0.277***	(0.104)
SUB_HY	-0.668**	(0.401)	InState	0.107	(0.159)
SUB_JA	0.131	(0.501)	Intl	-0.322*	(0.243)
SUB_MA	-1.120***	(0.486)	Needy	-0.076	(0.134)
SUB_MU	-1.303***	(0.457)	AcadRating	0.007	(0.023)
SUB_NS	-7.153	(14682754.049)	PctBotQ	1.720**	(1.003)
SUB_PC	-1.111***	(0.453)	PctTopQ	0.660	(0.699)
SUB_PH	-1.668***	(0.523)	ClassSize	0.011	(0.055)
SUB_PS	-0.856***	(0.398)	URM	-0.045	(0.205)
SUB_PY	-1.515***	(0.477)	Year1	-0.057	(0.349)
SUB_RE	-0.340	(0.478)	Year2	0.155	(0.262)
SUB_RS	0.906*	(0.580)	Year3	0.000	(0.000)
SUB_SO	-0.988***	(0.429)	Year4	0.244	(0.248)
SUB_SP	-0.490	(0.534)	Year5	-0.346	(0.369)
SUB_TH	-0.467	(0.608)	Sub_H	-0.702	(0.594)
N	715		Intercept	2.567**	(1.508)
Equation 2 : cutoffs			Div_N	-0.363	(1.389)
cutoff1	-1.097	(1.323)	Div_S	-0.270	(0.702)

* p<.2, ** p<.1, *** p<.05

Note: Professor dummy variables are omitted to save space.

Table B.2: 2 Selection Categories, 2 Stage Regression on Middle Achievers using Top/Bot 10%

Variable	Coefficient	(Std. Err.)	Variable	Coefficient	(Std. Err.)
Equation 1 : Selection 2 Categories			Equation 3 : Grade1		
Minority	-0.224*	(0.16)	_lambda	-0.698***	(0.23)
Female	-0.022	(0.09)	Minority	0.208**	(0.11)
InState	-0.058	(0.14)	Female	0.236***	(0.06)
Intl	0.151	(0.23)	InState	-0.044	(0.09)
Needy	-0.100	(0.11)	Intl	-0.155	(0.17)
AcadRating	0.014	(0.01)	Needy	-0.008	(0.08)
URM	0.082	(0.17)	AcadRating	0.022***	(0.01)
Demandc	-0.567***	(0.08)	PctBotQ	0.612	(0.54)
SUB_AH	-0.834***	(0.41)	PctTopQ	-0.741***	(0.33)
SUB_AN	-1.215***	(0.39)	ClassSize	-0.024*	(0.02)
SUB_BY	-0.263	(0.37)	URM	-0.250***	(0.11)
SUB_CH	-1.216***	(0.43)	Year1	0.000	(.)
SUB_CL	-0.293	(0.40)	Year2	0.391***	(0.11)
SUB_CN	0.072	(0.43)	Year3	0.249**	(0.13)
SUB_CO	-0.056	(0.43)	Year4	0.261***	(0.12)
SUB_DA	-0.234	(0.70)	Year5	0.188	(0.19)
SUB_DR	0.874*	(0.66)	Div_H	0.768***	(0.39)
SUB_ED	0.078	(0.49)	Div_N	0.568*	(0.38)
SUB_EN	-0.538*	(0.34)	Div_S	0.799***	(0.40)
SUB_ES	-0.303	(0.52)	Intercept	-2.469***	(0.88)
SUB_FS	-1.923***	(0.55)	Equation 4 : Grade2		
SUB_GR	0.027	(0.45)	_lambda	-0.068	(0.27)
SUB_GY	-0.793**	(0.41)	Minority	-0.066	(0.14)
SUB_HS	0.055	(0.41)	Female	0.155***	(0.08)
SUB_HY	-0.436	(0.34)	InState	0.070	(0.12)
SUB_JA	0.122	(0.43)	Intl	-0.529***	(0.18)
SUB_MA	-0.706**	(0.41)	Needy	0.021	(0.10)
SUB_MB	7.587	(981574.55)	AcadRating	0.023***	(0.01)
SUB_MU	-0.922***	(0.37)	PctBotQ	0.370	(0.78)
SUB_NS	-6.053	(669615.15)	PctTopQ	0.827**	(0.49)
SUB_PC	-0.975***	(0.39)	ClassSize	0.011	(0.04)
SUB_PH	-1.577***	(0.45)	URM	-0.025	(0.15)
SUB_PS	-0.696***	(0.34)	Year1	-0.097	(0.25)
SUB_PY	-1.529***	(0.44)	Year2	0.092	(0.20)
SUB_RE	-0.091	(0.42)	Year3	-0.167	(0.19)
SUB_RS	1.165***	(0.53)	Year4	0.000	(.)
SUB_SO	-0.917***	(0.36)	Year5	-0.442***	(0.22)
SUB_SP	-0.528	(0.42)	Div_H	-0.728*	(0.55)
SUB_TH	0.163	(0.50)	Div_N	-1.093**	(0.64)
			Div_S	-0.681	(0.62)
N	1060		Intercept	3.702***	(1.07)
Equation 2 : cutoffs					
cutoff1	-1.028*	(0.70)			

* p<.2, ** p<.1, *** p<.05

Note: Professor dummy variables are omitted to save space.

Table B.3: 2 Selection Categories, 2 Stage Regression on Middle Achievers using Top/Bot 5%

Variable	Coefficient	(Std. Err.)	Variable	Coefficient	(Std. Err.)
Equation 1 : Selection 2 Categories			Equation 3 : Grade1		
Minority	-0.168	(0.15)	_lambda	-0.644***	(0.23)
Female	0.051	(0.08)	Minority	0.200**	(0.10)
InState	-0.079	(0.13)	Female	0.187***	(0.06)
Intl	0.180	(0.21)	InState	-0.034	(0.09)
Needy	-0.075	(0.10)	Intl	-0.273**	(0.16)
AcadRating	0.004	(0.01)	Needy	-0.001	(0.07)
URM	0.009	(0.16)	AcadRating	0.024***	(0.01)
Demandc	-0.551***	(0.07)	PctBotQ	1.001**	(0.60)
SUB_AH	-0.946***	(0.40)	PctTopQ	-0.702**	(0.41)
SUB_AN	-1.257***	(0.38)	ClassSize	-0.019	(0.02)
SUB_BY	-0.399	(0.36)	URM	-0.295***	(0.11)
SUB_CH	-1.285***	(0.42)	Year1	0.000	(.)
SUB_CL	-0.254	(0.39)	Year2	0.440***	(0.11)
SUB_CN	0.045	(0.42)	Year3	0.255***	(0.13)
SUB_CO	0.080	(0.42)	Year4	0.241***	(0.11)
SUB_DA	-0.257	(0.70)	Year5	0.216	(0.18)
SUB_DR	0.811	(0.66)	Div_H	0.878***	(0.39)
SUB_ED	0.182	(0.48)	Div_N	0.544*	(0.37)
SUB_EN	-0.581**	(0.33)	Div_S	0.835***	(0.41)
SUB_ES	-0.472	(0.49)	Intercept	1.180*	(0.74)
SUB_FS	-1.986***	(0.54)	Equation 4 : Grade2		
SUB_GR	-0.045	(0.44)	_lambda	-0.286	(0.26)
SUB_GY	-0.865***	(0.39)	Minority	-0.087	(0.14)
SUB_HS	0.073	(0.39)	Female	0.156***	(0.07)
SUB_HY	-0.395	(0.34)	InState	0.059	(0.11)
SUB_JA	0.246	(0.42)	Intl	-0.586***	(0.18)
SUB_MA	-0.784***	(0.40)	Needy	0.039	(0.09)
SUB_MB	7.985	(2153349.47)	AcadRating	0.028***	(0.01)
SUB_MU	-0.964***	(0.36)	PctBotQ	-0.061	(0.70)
SUB_NS	-6.328	(1680637.32)	PctTopQ	0.198	(0.67)
SUB_PC	-1.039***	(0.38)	ClassSize	-0.032	(0.04)
SUB_PH	-1.321***	(0.41)	URM	0.043	(0.14)
SUB_PS	-0.633**	(0.33)	Year1	0.000	(.)
SUB_PY	-1.636***	(0.43)	Year2	0.176	(0.22)
SUB_RE	-0.225	(0.40)	Year3	-0.081	(0.22)
SUB_RS	1.215***	(0.52)	Year4	-0.019	(0.23)
SUB_SO	-0.893***	(0.35)	Year5	-0.040	(0.30)
SUB_SP	-0.485	(0.41)	Div_H	-0.193	(0.54)
SUB_TH	0.331	(0.48)	Div_N	-0.410	(0.62)
			Div_S	-0.075	(0.61)
N	1229		Intercept	2.859***	(1.14)
Equation 2 : cutoffs					
cutoff1	-1.404***	(0.59)			

* p<.2, ** p<.1, *** p<.05

Note: Professor dummy variables are omitted to save space.

Table B.4: 3 Selection Categories,
Selection Equation Output, Top/Bot 25%

Variable	Coefficient	(Std. Err.)
Minority	-0.275*	(0.19)
Female	-0.035	(0.11)
InState	0.084	(0.17)
Intl	0.276	(0.26)
Needy	-0.127	(0.14)
AcadRating	0.003	(0.02)
URM	0.044	(0.20)
Demand	0.065***	(0.01)
SUB_AH	-0.396	(0.41)
SUB_AN	-1.335***	(0.40)
SUB_BY	-0.282	(0.38)
SUB_CH	-2.011***	(0.51)
SUB_CL	-0.300	(0.41)
SUB_CN	-0.425	(0.46)
SUB_CO	0.322	(0.43)
SUB_DA	-1.259**	(0.65)
SUB_DR	0.121	(0.92)
SUB_ED	-0.457	(0.51)
SUB_EN	-0.863***	(0.34)
SUB_ES	0.050	(0.63)
SUB_FS	-2.453***	(0.60)
SUB_GR	-0.875***	(0.45)
SUB_GY	-1.490***	(0.44)
SUB_HS	0.083	(0.41)
SUB_HY	-0.632**	(0.35)
SUB_JA	-0.409	(0.44)
SUB_MA	-1.336***	(0.43)
SUB_MU	-1.387***	(0.41)
SUB_NS	-8.413	(747971.64)
SUB_PC	-1.261***	(0.40)
SUB_PH	-1.978***	(0.49)
SUB_PS	-1.053***	(0.36)
SUB_PY	-1.931***	(0.46)
SUB_RE	-0.287	(0.42)
SUB_RS	0.095	(0.48)
SUB_SO	-1.152***	(0.39)
SUB_SP	-0.340	(0.48)
SUB_TH	-1.067**	(0.56)
cutoffs		
cutoff1	-0.171	(1.22)
cutoff2	0.329	(1.22)

* p<.2, ** p<.1, *** p<.05

Table B.5: 3 Selection Categories,
Outcome Equation Output, Top/Bot 25%

	Grade1		Grade2		Grade3	
Variable	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**
_lambda	-0.431	(0.38)	0.700	(0.79)	1.360***	(0.64)
Minority	0.072	(0.12)	0.382	(0.33)	-0.583**	(0.32)
Female	0.252***	(0.07)	0.187	(0.19)	0.330***	(0.17)
InState	-0.062	(0.11)	0.410*	(0.28)	0.200	(0.29)
Intl	-0.030	(0.20)	-1.342***	(0.41)	0.114	(0.41)
Needy	0.009	(0.09)	0.019	(0.26)	-0.156	(0.24)
AcadRating	0.017	(0.02)	0.023	(0.04)	-0.002	(0.04)
PctBotQ	-0.314	(0.52)	1.248	(1.41)	1.215	(1.54)
PctTopQ	-0.840***	(0.39)	1.614**	(0.92)	-0.961	(1.16)
ClassSize	-0.004	(0.02)	-0.035	(0.08)	0.056	(0.08)
URM	-0.126	(0.12)	-0.842***	(0.34)	0.338	(0.33)
year_2011	0.000	(.)	0.000	(.)	0.000	(.)
year_2012	0.210*	(0.14)	0.727***	(0.32)	0.130	(0.80)
year_2013	0.236*	(0.16)	1.087***	(0.51)	-0.078	(0.76)
year_2014	0.147	(0.15)	1.262***	(0.49)	0.245	(0.77)
year_2015	0.108	(0.26)	1.251**	(0.74)	0.031	(0.87)
Div_H	0.750**	(0.43)	-2.021***	(0.96)	0.022	(1.34)
Div_N	0.375	(0.53)	-2.839***	(1.30)	-0.990	(1.22)
Div_S	0.545	(0.46)	-2.316***	(1.05)	0.989	(1.36)
Constant	2.004**	(1.06)	2.568	(2.51)	1.246	(2.49)
N	715					

* p<.2, ** p<.1, *** p<.05
Note: Professor dummy variables are omitted to save space.

Table B.6: 3 Selection Categories,
Selection Equation Output, Top/Bot 10%

Variable	Coefficient	(Std. Err.)
Minority	-0.095	(0.15)
Female	-0.077	(0.08)
InState	-0.072	(0.13)
Intl	0.183	(0.21)
Needy	-0.184**	(0.11)
AcadRating	0.009	(0.01)
URM	0.043	(0.16)
Demand	0.059***	(0.01)
SUB_AH	-0.621**	(0.36)
SUB_AN	-1.314***	(0.35)
SUB_BY	-0.341	(0.33)
SUB_CH	-1.674***	(0.40)
SUB_CL	-0.342	(0.36)
SUB_CN	-0.271	(0.39)
SUB_CO	0.518*	(0.38)
SUB_DA	-1.255**	(0.65)
SUB_DR	0.269	(0.61)
SUB_ED	-0.444	(0.44)
SUB_EN	-0.642***	(0.30)
SUB_ES	0.270	(0.44)
SUB_FS	-2.591***	(0.56)
SUB_GR	-0.773**	(0.41)
SUB_GY	-1.550***	(0.37)
SUB_HS	0.136	(0.36)
SUB_HY	-0.334	(0.31)
SUB_JA	-0.343	(0.38)
SUB_MA	-0.979***	(0.36)
SUB_MB	-0.155	(0.88)
SUB_MU	-1.029***	(0.33)
SUB_NS	-8.343	(669600.16)
SUB_PC	-1.242***	(0.34)
SUB_PH	-1.889***	(0.42)
SUB_PS	-0.997***	(0.31)
SUB_PY	-1.989***	(0.43)
SUB_RE	-0.035	(0.37)
SUB_RS	0.307	(0.42)
SUB_SO	-0.982***	(0.32)
SUB_SP	-0.529*	(0.38)
SUB_TH	-0.549	(0.45)
cutoffs		
cutoff1	0.181	(0.62)
cutoff2	0.682	(0.62)

* p<.2, ** p<.1, *** p<.05

Table B.7: 3 Selection Categories,
Outcome Equation Output, Top/Bot 10%

	Grade1		Grade2		Grade3	
Variable	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**
_lambda	-0.784***	(0.30)	-0.765**	(0.44)	1.038***	(0.46)
Minority	0.176*	(0.11)	0.270	(0.24)	-0.420***	(0.21)
Female	0.268***	(0.07)	0.118	(0.13)	0.196**	(0.12)
CO	-0.041	(0.10)	0.178	(0.18)	-0.080	(0.18)
Intl	-0.206	(0.17)	-1.067***	(0.33)	-0.400*	(0.29)
Needy	0.021	(0.08)	0.020	(0.17)	-0.041	(0.17)
AcadRating	0.026***	(0.01)	0.056***	(0.02)	0.015	(0.02)
PctBot10	0.615	(0.55)	0.112	(1.00)	-0.413	(1.14)
PctTop10	-0.705***	(0.33)	0.582	(0.71)	0.101	(0.61)
ClassSize	-0.013	(0.02)	-0.042	(0.05)	0.011	(0.05)
URM	-0.223**	(0.11)	-0.301	(0.25)	0.290*	(0.22)
year_2011	0.000	(.)	0.000	(.)	0.000	(.)
year_2012	0.407***	(0.12)	0.312*	(0.22)	0.125	(0.39)
year_2013	0.369***	(0.13)	0.579***	(0.29)	-0.107	(0.33)
year_2014	0.330***	(0.12)	0.153	(0.27)	0.184	(0.37)
year_2015	0.049	(0.21)	-0.273	(0.44)	-0.004	(0.55)
Div_H	0.984***	(0.39)	1.355	(1.39)	0.289	(0.41)
Div_N	0.882***	(0.44)	0.833	(1.54)	0.000	(.)
Div_S	1.022***	(0.41)	1.646	(1.59)	0.587	(0.51)
Constant	-3.332***	(0.89)	0.032	(1.10)	0.389	(1.55)
N	1060					

* p<.2, ** p<.1, *** p<.05
Note: Professor dummy variables are omitted to save space.

Table B.8: 3 Selection Categories, Selection Equation Output, Top/Bot 5%

Variable	Coefficient	(Std. Err.)
Minority	-0.097	(0.14)
Female	-0.010	(0.08)
InState	-0.077	(0.12)
Intl	0.152	(0.19)
Needy	-0.142*	(0.10)
AcadRating	0.002	(0.01)
URM	-0.012	(0.15)
Demand	0.061***	(0.01)
SUB_AH	-0.667**	(0.35)
SUB_AN	-1.408***	(0.34)
SUB_BY	-0.477*	(0.32)
SUB_CH	-1.738***	(0.39)
SUB_CL	-0.302	(0.34)
SUB_CN	-0.286	(0.38)
SUB_CO	0.549*	(0.37)
SUB_DA	-1.258**	(0.64)
SUB_DR	0.226	(0.61)
SUB_ED	-0.455	(0.43)
SUB_EN	-0.712***	(0.30)
SUB_ES	0.129	(0.43)
SUB_FS	-2.705***	(0.54)
SUB_GR	-0.860***	(0.40)
SUB_GY	-1.606***	(0.35)
SUB_HS	0.220	(0.34)
SUB_HY	-0.357	(0.30)
SUB_JA	-0.232	(0.37)
SUB_MA	-1.096***	(0.35)
SUB_MB	0.182	(0.79)
SUB_MU	-1.065***	(0.33)
SUB_NS	-8.711	(2759374.39)
SUB_PC	-1.329***	(0.33)
SUB_PH	-1.656***	(0.38)
SUB_PS	-1.012***	(0.30)
SUB_PY	-2.071***	(0.42)
SUB_RE	-0.206	(0.36)
SUB_RS	0.364	(0.41)
SUB_SO	-0.994***	(0.32)
SUB_SP	-0.433	(0.36)
SUB_TH	-0.476	(0.42)
cutoffs		
cutoff1	-0.150	(0.52)
cutoff2	0.350	(0.52)

* p<.2, ** p<.1, *** p<.05

Table B.9: 3 Selection Categories,
Outcome Equation Output, Top/Bot 5%

	Grade1		Grade2		Grade3	
Variable	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**
_lambda	-0.710***	(0.30)	-0.512**	(0.29)	0.782**	(0.41)
Minority	0.179**	(0.10)	0.262	(0.21)	-0.447***	(0.18)
Female	0.214***	(0.06)	0.224**	(0.11)	0.137*	(0.10)
InState	-0.031	(0.09)	0.113	(0.15)	-0.077	(0.16)
Intl	-0.286**	(0.16)	-1.109***	(0.26)	-0.422**	(0.24)
Needy	0.018	(0.07)	-0.138	(0.13)	0.056	(0.14)
AcadRating	0.025***	(0.01)	0.034***	(0.01)	0.022**	(0.01)
PctBot5	1.081**	(0.61)	-0.542	(0.88)	-0.631	(1.12)
PctTop5	-0.717**	(0.41)	-0.354	(0.99)	-0.748	(0.83)
ClassSize	-0.010	(0.02)	-0.040	(0.05)	-0.039	(0.05)
URM	-0.270***	(0.11)	-0.351**	(0.21)	0.381***	(0.19)
year_2011	0.000	(.)	-0.151	(0.26)	0.000	(.)
year_2012	0.456***	(0.11)	0.188	(0.23)	0.086	(0.36)
year_2013	0.360***	(0.12)	0.346*	(0.25)	-0.007	(0.30)
year_2014	0.308***	(0.11)	0.000	(.)	0.169	(0.33)
year_2015	0.096	(0.20)	-0.132	(0.35)	0.528	(0.51)
Div_H	1.086***	(0.40)	1.079**	(0.65)	-0.089	(0.38)
Div_N	0.824**	(0.44)	0.959	(0.99)	0.000	(.)
Div_S	1.056***	(0.42)	1.398**	(0.80)	0.105	(0.44)
Constant	-0.939	(0.98)	-2.973***	(1.25)	0.517	(1.43)
N	1229					

* p<.2, ** p<.1, *** p<.05
Note: Professor dummy variables are omitted to save space.

Table B.10: 4 Selection Categories,
Selection Equation Output, Top/Bot 5%

Variable	Coefficient	(Std. Err.)
Minority	-0.121	(0.14)
Female	-0.031	(0.08)
InState	-0.091	(0.11)
Intl	0.114	(0.18)
Needy	-0.150*	(0.09)
AcadRating	0.003	(0.01)
URM	0.014	(0.14)
Demand	0.050***	(0.01)
SUB_AH	-0.551*	(0.34)
SUB_AN	-1.307***	(0.33)
SUB_BY	-0.401*	(0.31)
SUB_CH	-1.656***	(0.38)
SUB_CL	-0.272	(0.33)
SUB_CN	-0.027	(0.37)
SUB_CO	0.601**	(0.35)
SUB_DA	-1.045**	(0.62)
SUB_DR	0.572	(0.59)
SUB_ED	-0.397	(0.41)
SUB_EN	-0.635***	(0.28)
SUB_ES	0.178	(0.41)
SUB_FS	-2.559***	(0.53)
SUB_GR	-0.711**	(0.39)
SUB_GY	-1.502***	(0.34)
SUB_HS	0.451*	(0.33)
SUB_HY	-0.361	(0.29)
SUB_JA	-0.227	(0.35)
SUB_MA	-1.014***	(0.34)
SUB_MB	-0.231	(0.66)
SUB_MU	-0.935***	(0.31)
SUB_NS	-8.371	(1022983.67)
SUB_PC	-1.256***	(0.32)
SUB_PH	-1.542***	(0.37)
SUB_PS	-0.943***	(0.29)
SUB_PY	-1.923***	(0.40)
SUB_RE	-0.071	(0.34)
SUB_RS	0.532*	(0.38)
SUB_SO	-0.902***	(0.30)
SUB_SP	-0.299	(0.35)
SUB_TH	-0.432	(0.40)
cutoffs		
cutoff1	-0.108	(0.50)
cutoff2	0.390	(0.50)
cutoff3	0.855**	(0.50)

* p<.2, ** p<.1, *** p<.05

Table B.11: 4 Selection Categories,
Outcome Equation Output, Top/Bot 5%

	Grade1		Grade2		Grade3		Grade4	
Variable	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**
_lambda	-0.635***	(0.29)	-0.502**	(0.28)	-0.118	(0.48)	0.313	(0.46)
Minority	0.182**	(0.10)	0.255	(0.21)	-0.062	(0.22)	-0.444***	(0.22)
Female	0.219***	(0.06)	0.232***	(0.11)	0.298***	(0.12)	0.053	(0.12)
InState	-0.028	(0.09)	0.116	(0.15)	0.142	(0.18)	-0.364**	(0.22)
Intl	-0.271**	(0.15)	-1.086***	(0.26)	-0.868***	(0.23)	-0.621**	(0.36)
Needy	0.016	(0.07)	-0.138	(0.13)	-0.113	(0.16)	0.316**	(0.18)
AcadRating	0.025***	(0.01)	0.034***	(0.01)	0.030***	(0.01)	0.020*	(0.02)
PctBot5	1.056**	(0.61)	-0.536	(0.87)	1.207	(1.58)	-0.070	(1.85)
PctTop5	-0.719**	(0.41)	-0.273	(0.99)	1.243	(1.43)	-1.939*	(1.29)
ClassSize	-0.010	(0.02)	-0.041	(0.05)	0.043	(0.07)	-0.020	(0.07)
URM	-0.278***	(0.11)	-0.345**	(0.21)	-0.116	(0.21)	0.285	(0.24)
year_2011	0.000	(.)	-0.128	(0.26)	0.000	(.)	0.000	(.)
year_2012	0.455***	(0.11)	0.208	(0.23)	0.257	(0.63)	-1.137***	(0.52)
year_2013	0.359***	(0.12)	0.360*	(0.25)	0.111	(0.39)	-1.229***	(0.48)
year_2014	0.308***	(0.11)	0.000	(.)	0.349	(0.48)	-0.924**	(0.52)
year_2015	0.155	(0.19)	-0.059	(0.32)	-0.930	(0.85)	0.116	(0.70)
Div_H	0.999***	(0.39)	1.052*	(0.64)	-0.304	(0.62)	-0.509	(0.58)
Div_N	0.715**	(0.42)	1.170	(0.97)	-0.259	(0.92)	0.000	(.)
Div_S	0.975***	(0.41)	1.347**	(0.78)	-0.548	(0.76)	-0.244	(0.67)
Constant	-0.764	(0.97)	2.425**	(1.25)	0.821	(1.19)	1.726	(1.72)
N	1229							

* p<.2, ** p<.1, *** p<.05
Note: Professor dummy variables are omitted to save space.

C OLS Results

The tables below display the results from the OLS foundational model (1). The overall results are the same, which suggests that selection bias is not affecting our model for the given sample.

Table C.1: OLS Model (1) on Middle Achievers

	Top/Bot 25%		Top/Bot 10%		Top/Bot 5%	
Variable	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**
Minority	-0.036	(0.11)	0.045	(0.09)	0.036	(0.08)
Female	0.258***	(0.07)	0.220***	(0.06)	0.223***	(0.05)
CO	0.027	(0.10)	-0.051	(0.08)	-0.048	(0.07)
Intl	-0.235	(0.25)	-0.358**	(0.20)	-0.417***	(0.18)
Needy	0.004	(0.08)	-0.022	(0.07)	-0.013	(0.06)
AcadRating	0.027**	(0.02)	0.032***	(0.01)	0.027***	(0.01)
PctBotQ	0.487	(0.59)	0.534	(0.48)	0.925**	(0.50)
PctTopQ	-0.273	(0.47)	-0.495**	(0.29)	-0.695**	(0.40)
ClassSize	-0.015	(0.02)	-0.005	(0.02)	-0.007	(0.02)
URM	-0.015	(0.11)	-0.096	(0.09)	-0.118*	(0.08)
year== 2011.0000	0.000	(.)	0.000	(.)	0.000	(.)
year== 2012.0000	0.115	(0.14)	0.223**	(0.11)	0.229***	(0.11)
year== 2013.0000	0.225*	(0.16)	0.288***	(0.13)	0.289***	(0.11)
year== 2014.0000	0.206*	(0.13)	0.337***	(0.11)	0.297***	(0.10)
year== 2015.0000	0.132	(0.19)	0.210*	(0.15)	0.262**	(0.15)
Division==H	0.199	(0.41)	0.276	(0.34)	0.250	(0.33)
Division==N	-0.286	(0.39)	-0.114	(0.30)	-0.064	(0.28)
Division==S	0.130	(0.42)	0.228	(0.35)	0.216	(0.32)
Constant	2.079***	(0.85)	-2.250***	(0.52)	1.709***	(0.48)
N	715		1060		1229	

* p<.2, ** p<.1, *** p<.05

Note: Professor dummy variables are omitted to save space. The column header represents the peer measures used.

Table C.2: OLS Model (1) on Low Achievers

	Top 25% Mid 50%		Top 10% Mid 80%		Top 5% Mid 90%	
Variable	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**
Minority	0.095	(0.17)	-0.275	(0.31)	-0.394	(0.32)
Female	0.237**	(0.13)	0.001	(0.18)	0.047	(0.29)
CO	-0.083	(0.16)	-0.158	(0.26)	-0.246	(0.29)
Intl	0.283	(0.50)	0.967**	(0.53)	-0.812***	(0.27)
Needy	0.046	(0.14)	-0.045	(0.24)	-0.435*	(0.32)
AcadRating	0.037***	(0.01)	0.054***	(0.02)	0.074**	(0.04)
PctTopQ	0.816	(0.89)	-3.448**	(2.00)	-0.347	(4.64)
PctMid50	0.769	(1.06)	-3.001**	(1.59)	-3.616	(2.82)
ClassSize	0.030	(0.04)	-0.104*	(0.08)	-0.228***	(0.05)
URM	-0.226	(0.20)	0.321	(0.32)	0.324	(0.36)
year== 2011.0000	0.517	(0.50)	-0.785*	(0.50)	-0.597	(0.80)
year== 2012.0000	0.679*	(0.49)	-0.689	(0.60)	-1.636*	(1.09)
year== 2013.0000	0.754*	(0.53)	-0.587	(0.62)	-0.952*	(0.57)
year== 2014.0000	0.915***	(0.44)	-0.630	(0.60)	-1.290*	(0.95)
year== 2015.0000	0.000	(.)	0.000	(.)	0.000	(.)
Division==H	1.139*	(0.74)	0.085	(0.51)	-1.036	(0.79)
Division==N	1.374***	(0.69)	0.913*	(0.70)	-0.079	(0.38)
Division==S	1.360*	(0.83)	0.000	(.)	0.000	(.)
Constant	-1.584	(2.05)	3.947**	(2.10)	7.271*	(5.33)
N	278		107		69	

* p<.2, ** p<.1, *** p<.05

Note: Professor dummy variables are omitted to save space. The column header represents the peer measures used.

Table C.3: OLS Model (1) on High Achievers

	Bot 25% Mid 50%		Bot 10% Mid 80%		Bot 5% Mid 90%	
Variable	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**
Minority	-0.060	(0.12)	-0.187*	(0.14)	-0.028	(0.20)
Female	0.056	(0.10)	-0.033	(0.13)	-0.074	(0.20)
CO	-0.045	(0.10)	0.042	(0.16)	0.140	(0.13)
Intl	-0.638***	(0.30)	-0.411	(0.46)	0.284	(0.31)
Needy	0.113	(0.10)	0.094	(0.13)	0.183	(0.27)
AcadRating	0.059***	(0.02)	0.059***	(0.03)	0.000	(0.05)
PctBotQ	0.191	(0.54)	-1.331	(1.48)	-1.305	(1.75)
PctMid50	1.395***	(0.47)	1.073**	(0.61)	1.080	(0.92)
ClassSize	-0.022	(0.02)	0.011	(0.04)	-0.008	(0.04)
URM	-0.120	(0.13)	0.065	(0.20)	-0.012	(0.30)
year== 2011.0000	0.000	(.)	-0.399*	(0.26)	0.119	(0.42)
year== 2012.0000	0.371***	(0.15)	0.000	(.)	0.141	(0.41)
year== 2013.0000	0.383***	(0.17)	-0.090	(0.27)	-0.025	(0.44)
year== 2014.0000	0.413***	(0.16)	0.002	(0.26)	0.000	(.)
year== 2015.0000	0.595***	(0.19)	-0.121	(0.35)	0.050	(0.36)
Division==H	0.411	(0.38)	0.422	(0.44)	0.352	(0.54)
Division==N	0.220	(0.38)	0.136	(0.45)	-0.068	(0.54)
Division==S	1.283***	(0.60)	1.493**	(0.84)	0.564	(0.70)
Constant	-1.411	(1.55)	-1.793	(2.24)	2.422	(2.93)
N	419		245		114	

* p<.2, ** p<.1, *** p<.05

Note: Professor dummy variables are omitted to save space. The column header represents the peer measures used.

D Robustness to Changes in Bidding System

In order to test the robustness of our results to the changes in the bidding system, we reran the regressions only on data from the years 2011-2014 (with the original bidding system). The overall results are the same, suggesting that our results are robust to the change in the bidding system.

Table D.1: Two Selection Categories

	Top/Bott~25% (Std. Err.)	Top/Bott~10% (Std. Err.)	Top/Bott~5% (Std. Err.)
Grade1			
PctBotQ	-0.234		
	(0.58)		
PctTopQ	-0.927***		
	(0.47)		
PctBot10		0.691	
		(0.59)	
PctTop10		-1.012***	
		(0.38)	
PctBot5			1.210**
			(0.65)
PctTop5			-1.076***
			(0.49)
Grade2			
PctBotQ	1.432		
	(1.20)		
PctTopQ	0.303		
	(0.87)		
PctBot10		0.402	
		(0.96)	
PctTop10		0.929**	
		(0.56)	
PctBot5			-0.218
			(0.76)
PctTop5			1.034*
			(0.74)

* p<.2, ** p<.1, *** p<.05
Note: Regressions run using 3 different peer measures and a two choice ordered probit for the first stage.

Table D.2: Three Selection Categories

	Top/Bott~25% (Std. Err.)	Top/Bott~10% (Std. Err.)	Top/Bott~5% (Std. Err.)
Grade1			
PctBotQ	-0.166		
	(0.58)		
PctTopQ	-0.880**		
	(0.47)		
PctBot10		0.676	
		(0.58)	
PctTop10		-0.991***	
		(0.38)	
PctBot5			1.202**
			(0.64)
PctTop5			-1.054***
			(0.49)
Grade2			
PctBotQ	-0.751		
	(1.13)		
PctTopQ	5.224***		
	(1.81)		
PctBot10		0.303	
		(1.08)	
PctTop10		-0.119	
		(0.73)	
PctBot5			0.136
			(0.85)
PctTop5			-0.181
			(1.01)
Grade3			
PctBotQ	0.959		
	(2.62)		
PctTopQ	-1.174		
	(1.34)		
PctBot10		-0.432	
		(1.50)	
PctTop10		0.786	
		(0.72)	
PctBot5			-1.362
			(1.27)
PctTop5			1.285*
			(0.86)

* p<.2, ** p<.1, *** p<.05

Note: Regressions run using 3 different peer measures and a three choice ordered probit for the first stage.

Table D.3: Four Selection Categories

	Grade1		Grade2		Grade3		Grade4	
	PctBot5	**PctTop5**	**PctBot5**	**PctTop5**	**PctBot5**	**PctTop5**	**PctBot5**	**PctTop5**
Top/Bott~25%	-0.156	-0.877**	-0.593	5.355***	-10.825***	-7.292**	0.000	5.958
(Std. Err.)	(0.58)	(0.47)	(1.08)	(1.77)	(1.80)	(3.94)	(.)	(4.93)
Top/Bott~10%	0.649	-0.984***	0.317	-0.100	0.491	0.781	-1.544	1.736
(Std. Err.)	(0.59)	(0.38)	(1.07)	(0.73)	(1.76)	(1.08)	(2.62)	(1.93)
Top/Bott~5%	1.192**	-1.046***	0.140	-0.150	-1.555	2.654	-3.668***	-0.542
(Std. Err.)	(0.64)	(0.49)	(0.84)	(1.02)	(1.50)	(2.23)	(1.84)	(1.12)

* $p<.2$, ** $p<.1$, *** $p<.05$
Note: Regressions run using 3 different peer measures and a three choice ordered probit for the first stage.