Ask and You Shall Receive? Gender Difference in Regrades in College

Hsueh-Hsiang Li*
Colorado State University

Basit Zafar †
Arizona State University

This version: December 31, 2018

Abstract

Women and men differ in their tendency to negotiate. This project examines whether male and female students experience different regrading rates in college. We analyze a unique administrative dataset that contains not only the final grade records but also any grade changes related to the records from a large 4-year public university. Our analysis based on the administrative records reveals that male students are 18.6 percent more likely than female students to receive favorable grade changes initiated by instructors. The gender difference in regrades persists across colleges and cannot be explained by observable characteristics of the students, instructors, and the classes.

*E-mail:.hsueh-hsiang.li@colostate.edu.
†E-mail: basitak@gmail.com.

1 Introduction

Women and men differ across a variety of behaviors, including their tendency to negotiate. The difference in propensity to negotiate contributes to a sizable portion of the gender gap in salaries and career advancement. This project examines whether such gender differences have consequences when individuals are in school. Specifically, we examine whether male and female students experience different success rates of grade changes in college. If men are more aggressive than women in bargaining for better grades (either on the extensive or intensive margin), they may be more likely to convince their instructors to alter their grades on the transcripts that are valued by many potential employers. Gender differences in willingness to ask and to negotiate may overstate the actual academic performance of male students and put equally capable female students at a relative disadvantage in the job market.

We analyze a unique administrative dataset that contains not only the final grade records but also any grade changes related to the records from a large 4-year public university, This administrative dataset documented different reasons of grade changes that allow us to distinguish changes that are resulted by student actions, university rules, or instructor initiations. Assuming that the distribution of grading errors is the same for both male and female students, we would expect to observe a similar grade correction pattern initiated by instructors for both male and female students. Our analysis based on the administrative records reveals that although women made up 53.4 percent of the grade records, they represented only 49.1 percent of the favorable grade changes initiated by instructors. The gender difference in students' grade changes persists across colleges and departments.

The difference in regrades by gender may have direct implications for labor market outcomes. Employers frequently require candidates who apply for entry-level positions to provide their transcripts, and many competitive positions require a minimum GPA. The difference in regrades by gender may push un-qualifying males above employers' GPA threshold. Prior studies primarily focus on the impact of gender differences in negotiations on the outcomes *after* individuals enter the labor force. There is little work on how gender differences in negotiations and willingness-to-ask *prior* to labor market entry may impact outcomes both before and after individuals enter the labor market. Our proposed project fills this gap by investigating whether the gender difference in propensity to bargain have implications for outcomes in college, and if so, whether that in turn has an impact on subsequent labor market outcomes.

2 Literature Review

Women are much less likely than men to initiate negotiations and attain less favorable outcomes when they do negotiate. For instance, Small et al. (2007) find that women were more likely than men to accept a low reward offer for their participation in the study without bargaining. In their study, only 2.5 percent of women demanded a higher payment for their participation - a stark contrast to 23 percent of men who made a similar request. Similarly, Babcock and Laschever (2009) find that among graduates from a prestigious MBA program, only 7 percent of the women negotiated their wage offers, while 57 percent of the men did. The gender difference in engaging in wage negotiation may contribute to the gender wage gap in the starting salary of the MBA graduates — 7.6 percent (approximtely $4,000) in this case. Leibbrandt and List (2014) conducted a field experiment by posting job advertisements and observing the negotiation behaviors by real job applicants. They find that when salaries were not explicitly made negotiable in the advertisement, men still negotiated for a higher wage, but women inclined to signal their willingness to accept a lower wage offer. However, the gender difference in wage negotiation disappears when salaries were posted as negotiable.

Researchers have also examined whether forcing women to negotiate can help narrow the gap. Exley et al. (2016) find that women attain worse returns from negotiation when it is mandatory than when it is optional. The research suggests that women choose not to negotiate because they expect the outcome to be unfavorable. Women's reluctance in initiating negotiations may stem from the backlash they believe they may face. For instance, Bowles et al. (2007) find that evaluators are more likely to penalize women than men for negotiating compensation. It is also widely acknowledged that women are more risk-averse than men in a variety of economic activities (Croson and Gneezy, 2009). If women are more likely than men to be penalized for engaging in negotiations and women are more risk-averse, we are likely to observe fewer regrade requests made by women.

The literature also finds evidence that the sex of negotiation partners affects individuals' propensity to negotiate and the bargaining outcomes. In ultimatum games, Solnick (2001) finds that women made more generous offers when paired with men than with women. Eckel and Grossman (2001) observed a similar pattern but they also find that women were most likely to reach an agreement when paired with other women in negotiation. In contrast, Sutter et al. (2009) find that when bargaining with a partner of the same sex, competition and retaliation intensified. Bowles et al. (2007) find that women were less inclined to initiate negotiations when the opponents were male. Dittrich et al. (2014) find that in a laboratory face-to-face bargaining game, women obtained worse negotiation outcomes than did men when they played the role of employees. The wages were higher when male employees negotiated with female employers than when female employees bargained with male employers. Hernandez-Arenaz and Iriberri (2018) analyzed field asymmetric bargaining games from a Spanish TV show and find that women respondents demanded a lower reward only when they negotiated with men although the opening offers did not vary by the sex of the proposers (the first mover in the games 95 percent of the time) and the respondents. To examine if the gender interactions affect the grade changes, we pay particular attention to the gender of both students and instructors.

The difference in regrades by gender may have direct implications for labor market outcomes. Employers frequently require candidates who apply for entry-level positions to provide their transcripts, and many competitive positions require a minimum GPA (Reshwan, 2016). The difference in regrades by gender may result in an inflation of the grades for male students and unfairly place female students in a disadvantage position. To the best of our knowledge, this is the first research to investigate the gender gap in grade changes and would help shed light on how the gender differences in negotiation patterns emerge before individuals enter the labor market. This pre-market gender difference is important because it has a direct implication on potential labor market outcomes for the college graduates.

3 Data

We analyze a unique administrative dataset from Colorado State University (CSU), a large 4-year public university. CSU was ranked 129th among all public and private universities nationwide in 2016.[1] Fall 2016 enrollment at CSU consisted of 23,768 on-campus undergraduate students. On average, the freshmen admitted in Fall 2016 had a 3.6 high-school GPA, a 25.2 ACT composite score (compared to 20.8 as the national average), and a 566.5 SAT critical thinking score and a 575.5 SAT math score (compared to 494 in critical thinking and 508 in mathematics as the national average).[2] Among these freshmen, approximate 56 percent were female and 25 percent were minorities.[3] Although the gender distribution of CSU freshmen was comparable to

[1] US News and World Report 2016

[2] The average CSU admission SAT score consists of a 574 SAT Math score and a 569 SAT Critical Reading score. The national average test scores were published on the ACT and SAT websites. The ACT Profile Report - National Graduating Class 2016 and 2016 College-Bound Seniors SAT Total Group Profile Report.

[3] Source: CSU The Fact Book 2016-17.

the national average (i.e. 56 percent), CSU was less ethnically diverse than the average U.S. university (i.e. 43 percent).[4]

This administrative dataset recorded not only the final grades but also any grade changes with different reasons. We focus primarily on the change reasons including grade entry errors, instructor corrections, and re-calculations - the only three options available to instructors when they submitted grade changes in the computing system. There were no clear instructions on the choices of the grade change codes and no verification mechanism in place to distinguish the assignment of regrade reasons. Therefore, instructors had the flexibility in assigning the grade change reasons among the three options. For this reason, we treat grade changes based on either of the three reasons as regrades by instructors. If the grade changes were based on students' own actions (e.g. taking a repeat-and-delete option by re-taking the same course to override the original grade with a new grade) or university rules (e.g. grades were automatically changed to F when students fail to meet the higher requirements for college writing and mathematics), they are not considered regrades made by instructors.

The administrative dataset contains 1,341,552 credited student-class records with letter grades from 64,857 students taught by 3,726 instructors during the years between 2010 and 2016. Excluding grade updates for the incomplete credits ("I") after students completed their work, there were 6,225 grade changes (0.46 percent) made by instructors during this time. Among the grade changes initiated by instructors, 94.6 percent (5,886 records) of the grades were corrected upward (i.e. when an initial grade was changed to a better grade). The overwhelming upward corrections among grade changes indicated that the risk of receiving a downward grade change was relatively small when students made re-grading requests. Although women made up 53.4 percent of the grade records, they represented only 49.2 percent of the

[4]Source: National Center for Education Statistics, Table 306.10. Total fall enrollment in degree-granting postsecondary institutions, by the level of enrollment, sex, attendance status, and race/ethnicity of student: Selected years, 1976 through 2016.

upward grade changes initiated by instructors. Table 1 shows the summary statistics by students' sex. Conditional on students' sex, the rate of upward grade changes initiated by instructors is 0.479 percent (2,991 records) for male students and 0.404 percent (2,895 records) for female students. Although grade changes were rare events, among these upward changes, the 0.075 percentage points difference represented that men were 18.6 percent more likely than women to receive an upward grade change by instructors. On the other hand, in the extremely sparse events (339 observations) of downward grade corrections (i.e. the initial grade was changed to a lower grade), male students (0.03 percent) were as 50-percent more likely than their female counterparts (0.02 percent) to receive such an adverse outcome.

Students showed additional differences in their course records by sex as presented in Table 1. For instance, female students took half of their classes from female instructors, while male students only took their classes from female instructors 38.9 percent of the time - potentially due to gender sorting into different college majors where the gender distribution of instructors is also uneven. For an average academic term, female students attained a higher average GPA by 0.2 points when compared with their male counterparts. The gender distribution of students also varies substantially across classes offered by different colleges. Women were more likely than men to take classes offered by the Colleges of Agriculture, Health and Human Sciences, Liberal Arts, Veterinary Medicine and Biomedical Sciences, and Intra-University. Nevertheless, the gender gap in grade changes was pervasive across colleges.

Columns 1 through 3 of Table 2 show that, except for the College of Agriculture and the Intra-University, women were less likely than men to receive an upward grade correction by instructors, and the gender difference was statistically significant in the Colleges of Business Administration, Liberal Arts, and Natural Sciences. In contrast, although male students were also more likely than female students to receive a downward correction of grades, the gender differences vary largely by colleges. Columns 4 through 6 of Table

2 shows that female students were more likely than male students to face the adverse outcomes in the Colleges of Agriculture, Business Administration, Natural Sciences, but male students were "punished" more in the College of Liberal Arts.

This preferential treatment of upward grade changes for male students is observed across both male and female instructors. Table 3 shows that female instructors changed grades for 0.412 percent of female students and 0.451 percent of male students, while male instructors changed grades for 0.408 percent of female students and 0.501 percent of male students. Although both male and female instructors seem to make grade changes for male students more often than for female students, this advantage enjoyed by male students was particularly pronounced among male instructors. Conditioning on student sex, female students receive similar treatments from both male and female instructors, while male students were more likely to receive upward grade changes from male instructors than from female instructors. The difference-in-difference measure captures the "men-helping-men" effect as 0.0545 percentage points.

Tables 4 shows that both female and male instructors were also more likely to revise the grade downward for male than for female students. Although both female and male instructors were more likely to change male students' grade to a lower grade, there was no significant gender interaction effect in terms of downward grade changes.

Students may have different propensity to request regrades based on the grade they originally received. Figure 1 shows that indeed students who received A+ or A ("A students") as their initial grade were very unlikely to experience a positive grade change because their grade already hit the upper bound. Other than these "A students", the positive grade changes do not demonstrate particular patterns depending on the initial grades. Students who received an initial grade of D had the largest probability (1.49 percent) of getting a boost in their final grade, followed by those with an initial grade of

C+ (0.92 percent), B- (0.9 percent), A- (0.79 percent) and F (0.72 percent). The relatively high frequencies observed from those receiving an initial grade of D or F were not surprising. Some departments (e.g. economics) require students to complete their courses at a grade of C and above in order to fulfill the major and/or minor requirements. Furthermore, if a student's average GPA falls below 2.0 (equivalent to a C), the student will be placed on academic probation and be dismissed if the probation has persisted for two semesters. However, the high frequencies of grade changes for those receiving A-, B-, and C+ also indicate that the upward grade changes were not limited to the failing students.

Students did not always argue for a better grade simply because they performed poorly in the class. Frequently, students were motivated to make such an argument when the grade they received was below their expectations. We calculate the difference between the students' average GPA during the academic term and the initial grade they received from the class to capture the expectation gap and plot the density of upward grade changes against such an expectation gap. Figure 2 shows that the upward grade changes were near zero when the expectation gap was negative, i.e. when the student's performance in a given class was better than his average performance in other classes during the same semester. However, the propensity to receive an upward grade correction was increasing in the expectation gap when the student's class performance was below his average grade in the same semester.

Conditional on grade records that were changed by instructors, the majority of the changes were corrections by moving up one letter grade. Table 5 documents the matrix of the grade transitions. The upward change by one letter grade include changes from: A to A+ (1.19 percent), A- to A (11.73 percent), B+ to A- (5.48 percent), B to B+ (4.47 percent), B to A (13.98 percent), B- to B (6.63 percent), C+ to B- (3.08 percent), C to C+ (1.82 percent), C to B (10.70 percent), D to C (9.69 percent), and F to D (2.41

percent).[5] The upward one-letter grade changes account for 71 percent of total grade changes (including both the positive and negative grade changes). The intensity of the grade changes at the margin implies that instructors may be sympathetic towards students when their grade was close to the margin between two letter grades.

The frequencies of upward grade change received by students also vary by their sex. Figure 3 presents the conditional distribution of upward grade changes by student sex. Conditional on upward grade changes, most of the students received a single upward grade change during this period. Among female students, 92.7 percent among those who receive upward grade changes were awarded the upward grade correction once. For male students, 92 percent were this case. The separate density plots by student sex exhibit that the male distribution first order stochastic dominates the female distribution, i.e. women were much less likely than men to receive multiple upward grade changes.

4 Empirical Specifications and Analysis

To examine whether gender differences are present in grade changes among college students, we analyze our data with the following two empirical specifications:

$$Y_{ij} = \alpha_0 + \alpha_1 Male_i + \alpha_2 X_i + \alpha_3 Z_j + \epsilon_{ij}, \tag{1}$$

$$\begin{aligned} Y_{ij} =& \beta_0 + \beta_1 Female_i Male_j + \beta_2 Male_i Female_j + \beta_3 Male_i Male_j \\ &+ \beta_4 X_i + \beta_5 Z_j + \eta_{ij}, \end{aligned} \tag{2}$$

[5] Grade changes from B to A or from C to B are frequently an adjustment of one letter grade when plus and minus scales were not used.

In both equations, Y_{ij} is a binary variable and assumes the value of one when the student i received a grade change (e.g. upward grade change or downward grade change) in class j and zero for those with no changes. X_i are characteristics of student i, such as student's class standing (i.e. freshman, sophomore, junior, and senior), GPA, and grade, and Z_j captures class-specific information, such as instructor's position (i.e. tenured track faculty at different levels, instructors, graduate teaching assistants), department, and colleges. $Male_i$ in equation 1 is an indicator for male students and α_1 captures any gender difference in the regrades. Equation 2 investigates whether the grade change patterns vary by the gender interactions between students and instructors. In equation 2, β_0 capture the average grade change rate for the omitted reference gender pair group, i.e. female student and female instructor. The variables $Female_i Male_j$, $Male_i Female_j$, and $Male_i Male_j$ are indicators for the gender pairs of female student and male instructor, male student and female instructor, and male student and male instructor, respectively. If β_1, β_2, and β_3 are simultaneously indistinguishable from zero, then the assumption of any gender interaction effects would be rejected.

4.1 Upward Grade Changes

We first present the analysis of equation 1 for upward grade changes in Panel A of Table 6. The raw gender difference in upward grade change is 0.0749 percentage points shown in Column 1. Adding controls for the colleges in which the courses were offered does not reduce the gender gap but instead increases the gap slightly to 0.0767 percent points favoring male students (Column 2). Column 3 controls for the departments where the courses were listed, and it reduces the gender difference to 0.0702 percentage points. The different grade change rates by departments explain approximately 6.3 percent of the gender gap in upward regrades.

Adding additional controls for instructors' positions (tenured tracked assistant professor, tenured tracked associate professor, tenure-tracked full professor, non-tenure tracked instructors, graduate teaching assistants) in Column 4 and students' class standing (freshmen, sophomores, juniors, seniors) in Column 5 of Panel A results in little change to the gender gap.

As Figure 2 shows that students with an average GPA higher than the grade they received from a course were more likely to receive an upward grade correction by instructors, we included the measure of the grade expectation gap (i.e. GPA − Grade) as an additional control in Column 6 of Panel A. Once the grade expectation gap is controlled for, the male advantage reduces to 0.0706 percentage points. In other words, the differences of the grade expectation gap explain 5.8 percent of the gender difference in upward regrades. This specification $Y_{ij} = \alpha_0 + \alpha_1 Male_i + \alpha_2(GPA - Grade) + \epsilon_{ij}$ implicitly assumes that the coefficient on $Grade$ is equal to the negative coefficient on GPA. Alternatively, if $Y_{ij} = \alpha_0 + \alpha_1 Male_i + \alpha_2 GPA - \gamma Grade + \epsilon_{ij}$ is estimated, the effects of GPA and course grade on the propensity of grade changes are allowed to vary. Indeed, the F-test rejects the hypothesis that $\alpha_2 = \gamma$ (result omitted from Table 6). To accommodate the non-linearity of the letter grades, Column 7 includes both the students' average GPA and separate indicators for the letter grades that students received in the class during the same academic semester as additional controls. Both the GPA and the class grades significantly influence the likelihood of upward grade changes for students, but they failed to explain the gender gap in upward regrades.

Prior literature documented important gender dynamics between the two parties engaging in negotiation. To investigate the effect of gender interactions between students and instructors, Panel B of Table 6 estimates Equation 2 and tests whether the coefficients on all other gender pairs (i.e. Female student x Male instructor, Male student x Female instructor, and Male student x Male instructor) compared to the reference group (Female student x Female instructor) are jointly zeros. Throughout all specifications, we find strong evidence of gender interaction effects and the male advantage is particulary pronounced when the male students are in the classes taught by male instructors.

Including all the forth-mentioned control variables in Table 6, Column 1 of Panel A in Table 7 reports the baseline estimate of a 0.079 percentage points male advantage in upward regrades. The marginal effects are insensitive to the adoption of the alternative Probit and Logit models, and the male advantage in upward grade changes is 0.082 and 0.085 percentage points, respectively from these two models (Columns 2 and 3 of Panel A). When we take into an account the gender interactions between students and instructors, Panel B of Table 7 shows that female students may face a relative disadvantage in interacting with male instructors. The probability of receiving an upward grade change reduces by 0.046 percentage points in the Probit model and 0.043 percentage points in the Logit model when female students take the class from a male instructor than from a female instructor (Row 1 of Panel B). Preferential treatments for male students in upward grade corrections have similar magnitudes across all these different models. Male students who take classes from female instructors are 0.038–0.044 percentage points more likely than female students who take classes from female instructor to receive an upward grade change. Among all these groups, male students who enroll in male instructors' classes are the most advantageous in term of grade changes — the likelihood of them receiving an upward grade change is 0.063 - 0.075 percentage points higher than female students who enrolled in female instructors' classes.

To identify the gender interaction effects between students and instructors, it requires information on the sex of both the students and the instructors. Although information on students is complete, there are approximately 6.4 percent of the grade records with missing information on the instructor's sex. If the data are not missing at random, the estimates may be biased. To address this data issue, we conducted a partial identification analysis to estimate the upper bound and lower bound of the point estimate of the linear probability model. The bound analysis assumes that instructors with missing sex information are either all male or all female. Table 8 reports the estimates of upper bounds and lower bounds. The male advantage remained statistically different from zero for both upper and lower bounds. For the most advantageous group (i.e. male students in male instructors' classes), the initial point estimate indicates that they have 0.075 percentage point advantage compared with female students in female instructors' classes. Assuming all the missing values are from female instructors, this advantage could be as high as 0.087 percentage points. In contrast, if all the missing values are from male instructors, this advantage reduces to 0.049 percentage points but remains significant.

Compared to the baseline linear probability model in Column 1 of Panel A in Table 9 with a 0.079 percentage points male advantage in upward regrades, we explore the source of the gender difference. To investigate the sensitivity of the estimate, we first removed students who received multiple upward regrades from the analysis. We find that the gender gap reduces but remains substantial at 0.065 percentage points in Column 2 of Panel A. This result suggests that approximately 18 percent of the male advantage was due to the fact that male students were more likely to receive multiple upward grade changes than were their female counterparts during their time in college.

If the instructors who made frequent grade adjustments tend to be teaching classes taken predominantly by male students, we might over-estimate the gender difference in the regrade outcomes. In the third column of Panel A in Table 9, we controlled for instructor fixed effects, and the gender gap decreases to 0.0677 percentage points but remains significant.

Next, we consider that students whose grade hit the upper bound (A+ and A) may not *need* to request grade changes. We removed them from the analysis in Column 4 of Panel A in Table 9. The gender gap actually increases to 0.1135 percentage points because female students are over-represented in this grade group. On the other end of the grade spectrum, students who hit the lower bound (F) may have greater incentives to request regrades because they face no downside risk. By removing students who received a failing grade, the gender gap becomes 0.081 percentage points — similar to the magnitude in the baseline model.

Panel B of Table 9 examines the sensitivity of gender interaction effects. In the baseline model, compared with the omitted gender pair $Female_iFemale_j$, the coefficient on $Female_iMale_j$ is small (-0.000263) and insignificant - indicating that male and female instructors treated female students with a similar standard in regrades. However, male students are better off than female students in terms of regrades. Compared with female students in female instructors' class ($Female_iFemle_j$), male students in female instructors' class ($Male_iFemale_j$) enjoy a 0.044 percentage points advantage in regrades, and this advantage for male students in male instructors' class ($Male_iMale_j$) is even larger at 0.075 percentage points. Although the coefficients change to some degree across the sensitivity analyses, the different effects by gender pairs persist. Note that in Column 3 when instructors fixed efffects are controlled for, only two gender interaction effects can be separately identified. Therefore, the coefficients on $Male_iFemale_j$ and $Male_iMale_j$ are compared with the omitted group of female students, regardless of the gender of the instructors. Similarly, the male advantage in regrade is present and most salient when the male students are taking the courses from male instructors.

4.2 Downward Grade Changes

We analyze the same equations with the downward regrade outcomes in Table 10. As stated earlier, male students were also more likely (by 0.0103 percentage points as reported in Column 1 of Panel A) than female students to receive a downward regrade. Although the result is insensitive to the inclusion of colleges, instructors' positions, students' class standing, GPA, and initial grades, the gender difference disappears when we control for the departments of class offerings. This result suggests that some departments were more likely to change students' grade to a lower grade than were other departments, and male students were more likely to take classes from these departments.

When we control for all these forth-mentioned independent variables, the gender difference in downward regrades becomes trivial (0.0002 percentage points) and insignificant as presented in the baseline model in Column 1 of Panel A in Table 11. The absence of gender difference in downward regrades is persistent across the sensitivity analyses when we remove students who received multiple negative regrades, control for instructor fixed effects, remove students with an initial grade of A+ or A, and remove students with an initial grade of F (Columns 2 through 5 of Panel A). When we turn our attention to potential gender interaction effects in Panel B, we also find no evidence to support any interactive gender dynamics between students and instructors with an only exception case when we remove students with an initial grade of A+ or A from the analysis. Compared with all other gender pairs, female students with an initial grade below A and took classes from female instructors had the lowest chance of getting a lower grade as a regrade result. In other words, facing students with an initial grade below A, female instructors were more lenient toward female students than towards male students. Male instructors treated equally both female and male students with a grade below A.[6]

4.3 Effect on Grades

Table 12 shows that conditional on the initial grades, students with regrades had average grades that were comparable to those of the students without regrades. However, after the regrades, the grade improvement for these students were ranging from 0.15 for the "A students" to 1.96 for the "F students." The increment is particularly large when the student's initial grade was low because some plus and minus grades were unavailable at the lower

[6]The F-test of $H_0 : \beta_{FM} = \beta_{MM}$ has a p-value of 0.5784.

end of the grade distribution, and the upward change would elevate the grade points substantially.[7]

To analyze the gains from regrades for female and male students separately, Tables 13 and 14 repeat the same exercise of the grade analysis. Interestingly, conditional on the initial grade received by students who experienced a regrade, female students gained a larger improvement in their grade than did male students on average. Hence, although female students were less likely to receive an upward grade change made by instructors, they were more effective in achieving a higher final grade. This finding is consistent with Exley et al. (2016) that women negotiated only when they knew they would do well. This might be the case if female students only asked for regrades when apparent errors were involved in the initial grade assigned to them, and they received the change in grade for what they should have gotten in the first place.

5 Instructor Survey

The gender difference in grade changes may be a result of three distinct scenarios: 1) male students are more likely than female students to ask instructors for grade changes although instructors treat all requests equally; 2) the propensity to ask is the same for both male and female students, but the outcomes are more favorable for males when they ask; and 3) female students make regrading requests *during* the semester which in turn lowers their demand for regrading requests at the *end* of the semester. Because the administrative transcript records do not contain any information that allows us to understand which scenario accounts for the unequal outcomes, we conducted a survey on instructors to elicit their recollection on grade change requests made by students from the past semesters. The survey allows us

[7]For instance, C-, D+, and D- are not eligible options for the final grade entries at CSU.

to examine if male students are actually more likely than female students to make grade changes requests.

One hundred and fifty-four (154) instructors who experienced regrade requests in previous semesters completed the survey. The respondents are slightly over-represented by female instructors (58 percent female in the survey compared to 48 percent in the administrative records) and under-represented by non-tenure tracked instructors (26 percent non-tenure tracked instructors in the survey as oppose to 49 percent in the administrative records). Weighted by class sizes, male students represent 52.5 percent of the classes in the sample - higher than the 46.6 percent in the administrative records. Since the survey focuses on instructors who had experienced regrade requests in the past, it implies that classes with more male students are associated with an increase in the probability of regrade requests. Instructors reported that 5.94 percent of their students requested to change their final grade at the end of the semester, and 11.2 percent of students requested regrades during the semester. Weighted by class sizes, instructors who had experienced regrade requests reported that 0.727 percent of the grades were corrected to a better grade at the end of the semester. The reported upward corrections were more frequent than the actual records at the 0.439 percent because the instructor survey only elicited participation of instructors who has experienced regrade requests in the past.[8]

Figure 4 reports the gender distribution in class and among regrade requests. Male students made up 57.1 percent of regrade requests at the *end* of the semester – a larger proportion than their representation in class (52.5

[8]In the sub-sample collected between November 30 and December 9, 2018, we allowed all instructors to participate. The reported upward regrades comprises 0.354 percent of the students reported by instructors in the sub-sample - close to the statistic from the administrative records. This exercise reconfirms that the difference between the instructor survey and the administrative records is attributed to the exclusion of instructors who never experienced regrade requests in the survey. To keep the sample consistent, we report all the results conditional on instructors who had experienced some regrade requests at the end of the semester in the past.

percent). The gender differential outcome at the end of the semester could be a result of differences in the timing of regrade requests by student sex. However, the survey shows that male students were still over-represented in the regrade requests *during* the semester (58 percent of the requests versus 52.5 percent of their representation in class). This result rejects the potential scenario that female students manage to get their grade corrected before the semester ends to avoid the need to request a better grade when the final grade is assigned.

Using an alternative measure by comparing the fraction of males requesting regrades at the end of (during) the semester and the fraction of males in class, we find that 49.3 percent (55 percent) of the instructors (weighted by class sizes) reported more male students among regrade requests than their representation in class, 35.6 percent (29.1 percent) reported fewer male students among regrade requests than their representation in the class, and 15.1 percent (15.9 percent) reported the exact representation of male students in regrade requests as well as in the class. The odds-ratio of over-representation to under-representation in regrade requests is 1.39 (1.89) for male students and 0.72 (0.53) for female students at the end of (during) the semester.

Additionally, when asked whether male or female students were more aggressive in requesting regrades, 46.7 percent (35.1 percent) of instructors (weighted by students taught) indicated that male students were more aggressive than female students, only 10.3 percent (14.6 percent) of instructors (weighted by students taught) expressed that female students were more aggressive than male students, and the remaining stated that male and female students were similarly aggressive at the end of (during) the semester.

With regards to the outcomes of regrade requests, Figures 5 and 6 show that the change patterns conditional on student requests were indistinguishable for male and female students both during and at the end of the semester. Majority (76 percent) of instructors (weighted by class sizes) further reported an impression that male and female students are equally successful in regrades. Nonetheless, 21.8 percent of instructors (weighted by class sizes) reported that female students are more successful than male students in regrades, while only 2.27 percent of instructors (weighted by class sizes) reported that male students are more successful than female students in regrade outcomes. These results from the instructor survey suggest that most instructors treated regrade requests by male and female students equally. If there is any difference, the advantage is more likely to be on the female students' side.

Instructors also provide assessment of the ability ranking for the average student who made regrade requests at the end of the semester, during the semester, and in the class by student sex. Figure 7 shows that female students were ranked higher than male students in their overall ability when compared with male students in class. Among students who made regrade requests, female students were also ranked higher than male students in the ability distribution. This finding is consistent with the pattern of regrades records from the administrative dataset where male students dominate the low-end of the grade distribution and female students over-represent the high-end of the grade distribution among regrades.

The instructor survey results provide evidence that the most likely scenario is that male students are simply more likely than female students to ask for regrades and the request patterns persist throughout the semester. Hence, even if instructors are granting male and female students regrades at the same rate, the outcome may still favor male students simply because they ask more frequently.

6 Conclusion

We use a unique administrative data set to analyze the regrades in a large public university and find that male students were 18.6 percent more likely than female students to receive an upward grade correction granted by instructors. This male advantage in regrades is most salient when male students were interacting with male instructors. The gender difference can hardly be explained by observable characteristics of the class, instructors, and students. The survey results from instructors provide suggestive evidence that male students are dis-proportionally more likely to make regrade requests which result in favorable outcomes for men. This gender gap may place equally capable male and female students in an unequal footing before they even enter the labor market. If employers rely on college transcripts as a screening mechanism, this gender gap may contribute to the unfavorable treatment to women in the labor market. Our findings call for attention to the college grading practices. Further investigation on the sources of such persistent gender bias would help shed light on the mechanism of the gender differences and inform potential policy interventions.

References

Babcock, L. and S. Laschever (2009). *Women don't ask: Negotiation and the gender divide*. Princeton University Press.

Bowles, H. R., L. Babcock, and L. Lai (2007). Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask. *Organizational Behavior and human decision Processes 103*(1), 84–103.

Croson, R. and U. Gneezy (2009). Gender differences in preferences. *Journal of Economic literature 47*(2), 448–74.

Dittrich, M., A. Knabe, and K. Leipold (2014). Gender differences in experimental wage negotiations. *Economic Inquiry 52*(2), 862–873.

Eckel, C. C. and P. J. Grossman (2001). Chivalry and solidarity in ultimatum games. *Economic Inquiry 39*(2), 171–188.

Exley, C. L., M. Niederle, and L. Vesterlund (2016). Knowing when to ask: The cost of leaning in. Technical report, National Bureau of Economic Research.

Hernandez-Arenaz, I. and N. Iriberri (2018). Gender differences in alternating-offer bargaining: An experimental study. *history 15499*, 2.

Leibbrandt, A. and J. A. List (2014). Do women avoid salary negotiations? evidence from a large-scale natural field experiment. *Management Science 61*(9), 2016–2024.

Reshwan, R. (2016, April). Does gpa matter when applying for a job?

Small, D. A., M. Gelfand, L. Babcock, and H. Gettman (2007). Who goes to the bargaining table? the influence of gender and framing on the initiation of negotiation. *Journal of personality and social psychology 93*(4), 600.

Solnick, S. J. (2001). Gender differences in the ultimatum game. *Economic Inquiry 39*(2), 189–200.

Sutter, M., R. Bosman, M. G. Kocher, and F. van Winden (2009). Gender pairing and bargaining—beware the same sex! *Experimental Economics 12*(3), 318–331.

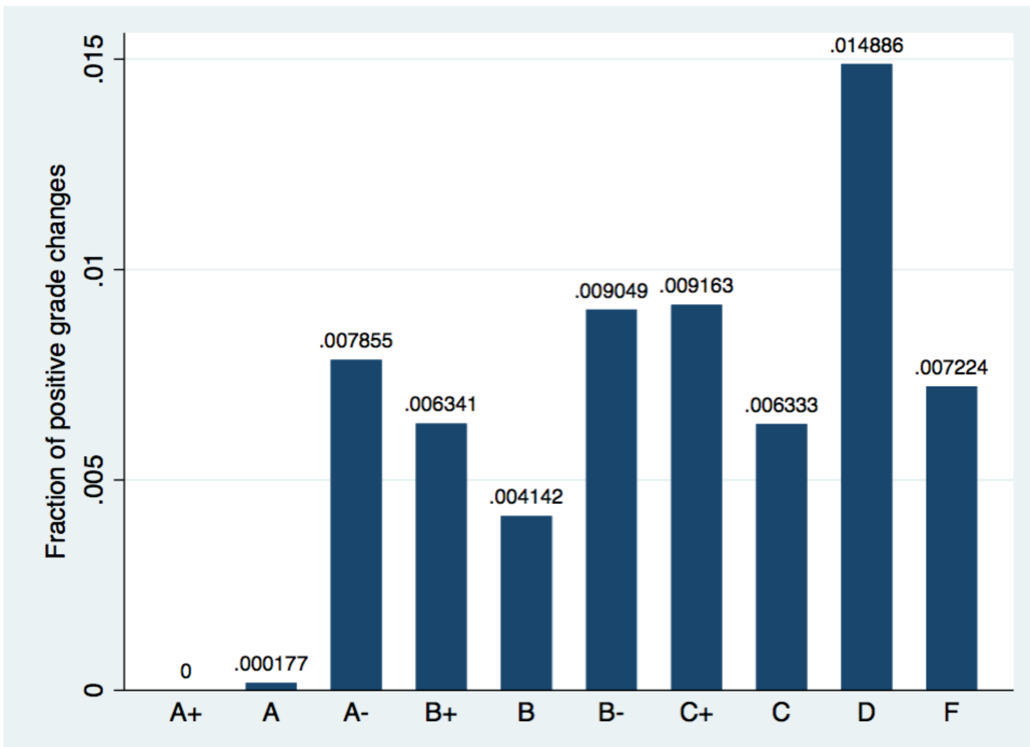

Figure 1: Fraction upward grade changes conditional on initial grades

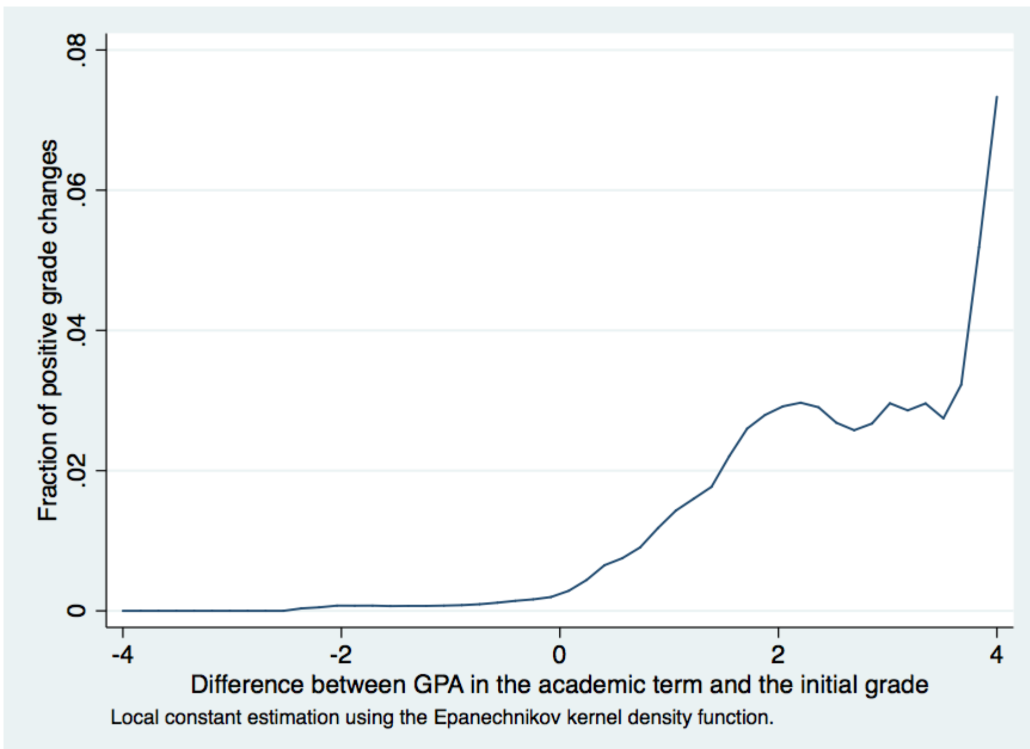

Figure 2: Non-parametric estimate of propensity of upward grade changes by the grade expectation gap

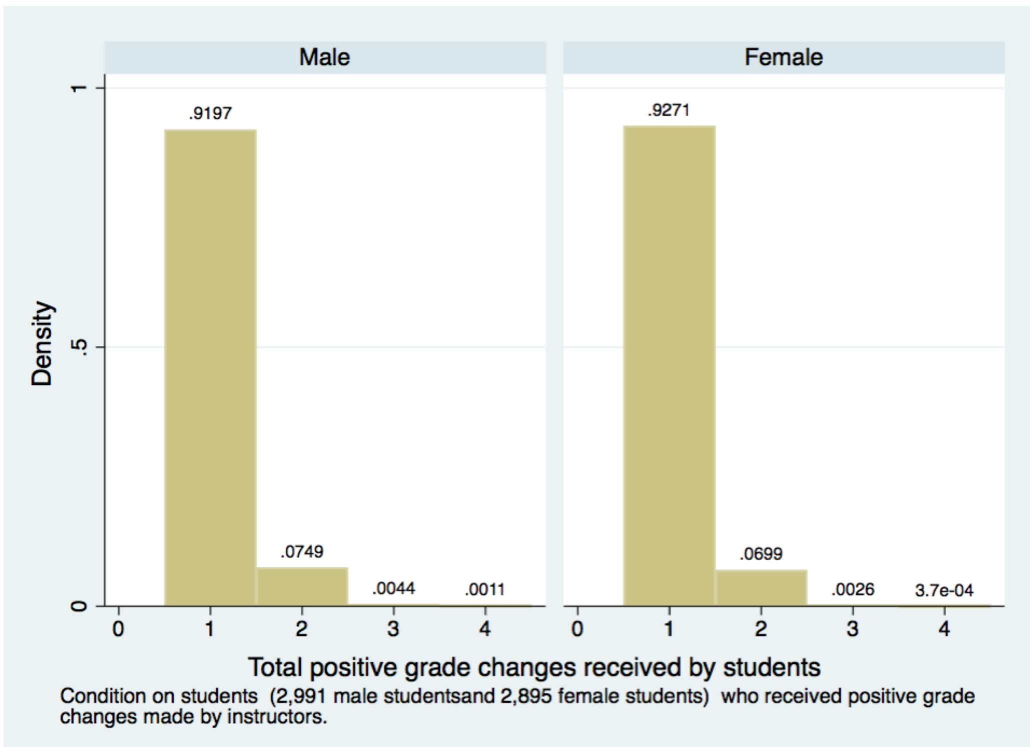

Figure 3: Conditional distribution of grade changes by student sex

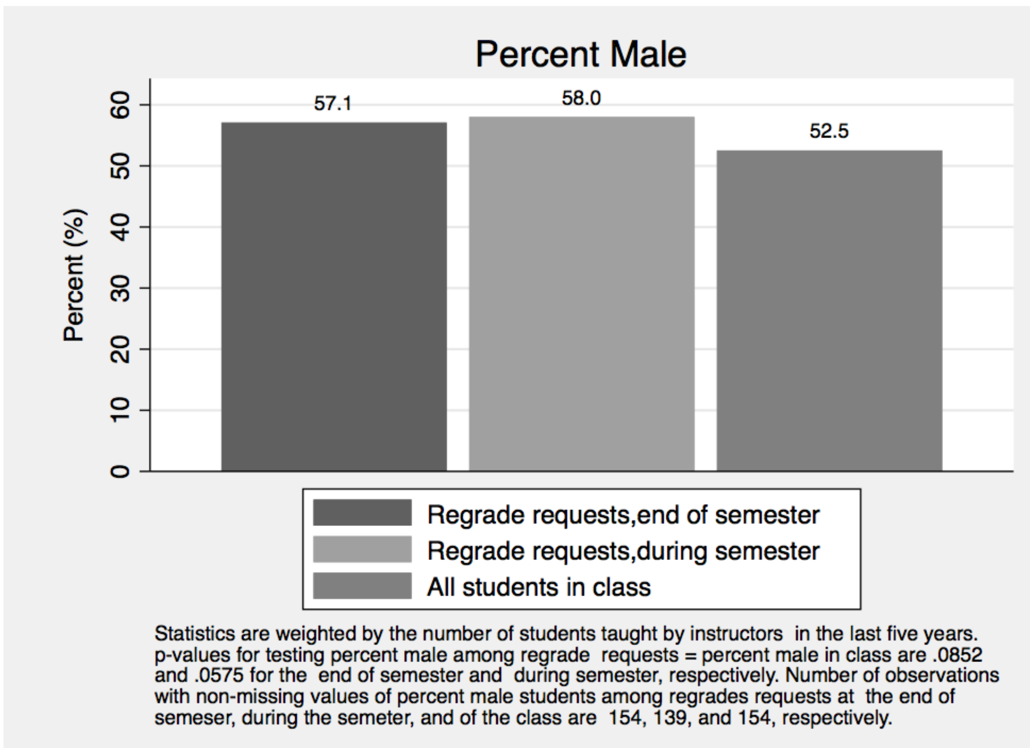

Figure 4: Instructor survey: Percent male students among regrade requests and in the class

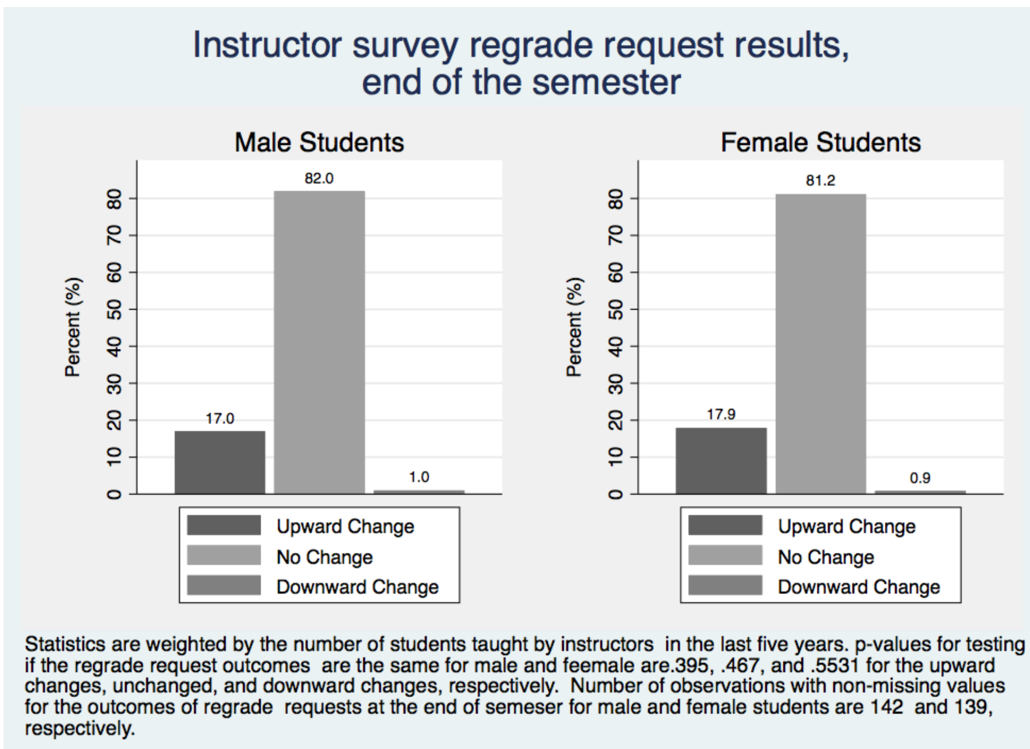

Figure 5: Instructor survey: regrade results by student sex at the end of semester

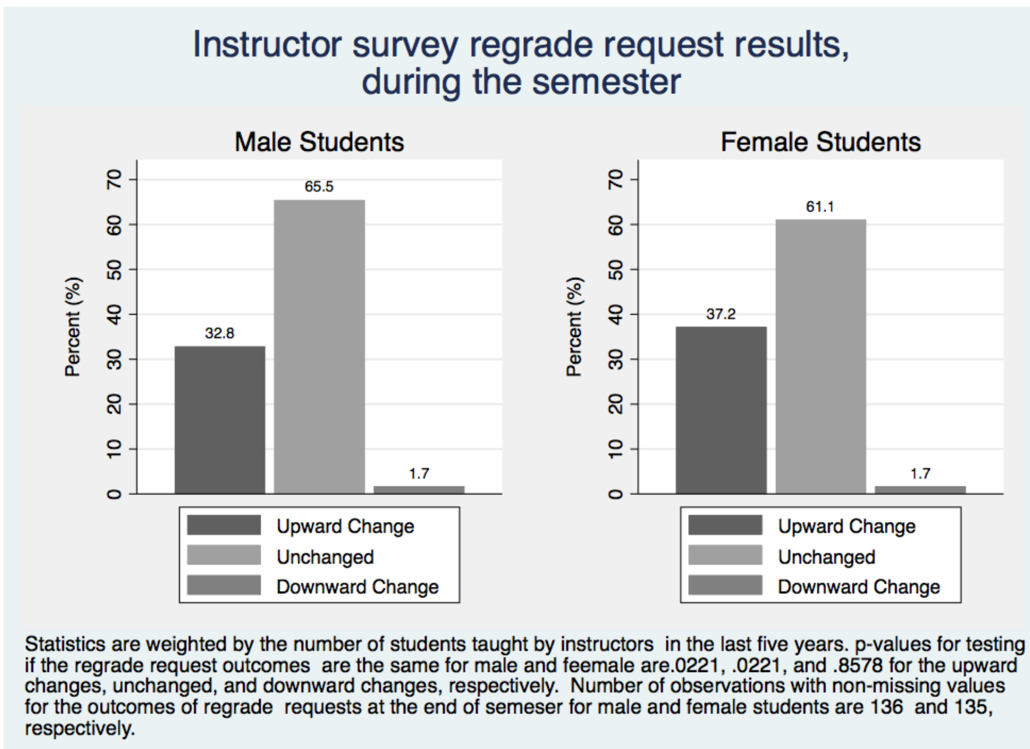

Figure 6: Instructor survey: regrade results by student sex during semester

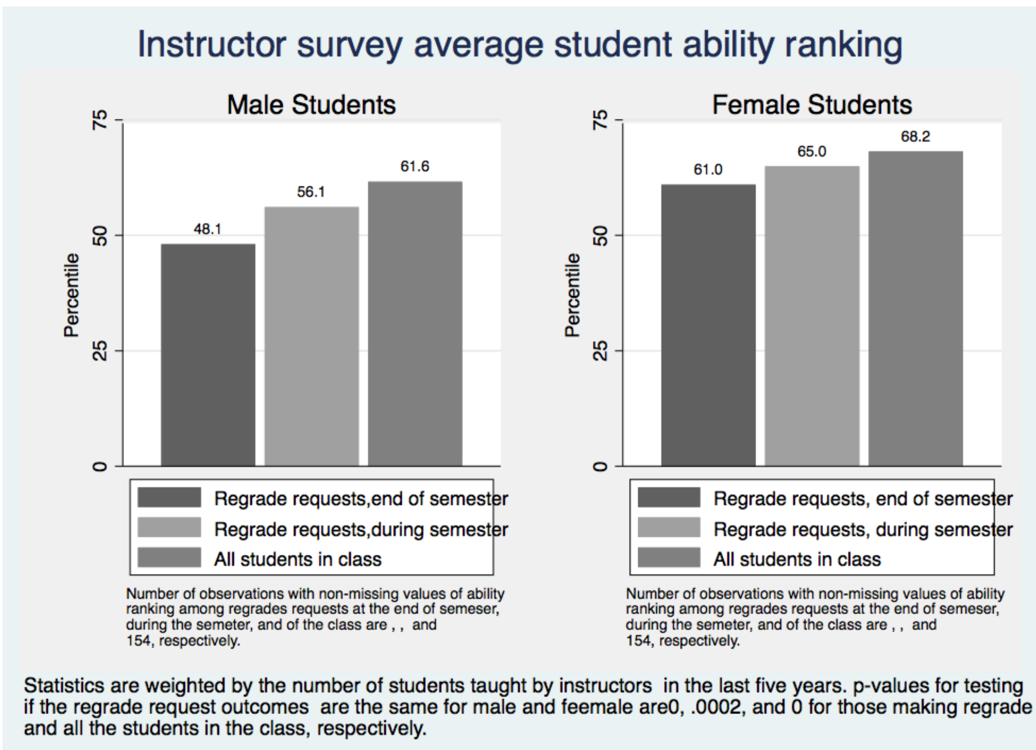

Figure 7: Instructor survey: student ability ranking by student sex

Table 1: Sumamry statistics by student sex

	Female	Male	Difference
Grade change	0.00424	0.00509	-0.00085***
	(0.06501)	(0.07119)	(0.00012)
	[716,772]	[624,780]	
Positive grade change	0.00404	0.00479	-0.00075***
	(0.06343)	(0.06904)	(0.00011)
	[716,625]	[624,588]	
Negative grade change	0.00021	0.00031	-0.00010***
	(0.01435)	(0.01757)	(0.00003)
	[713,877]	[621,789]	
Female instructor	0.50026	0.38942	0.11084***
	(0.50000)	(0.48762)	(0.00088)
	[671,276]	[585,026]	
Term GPA	3.14551	2.94470	0.20081***
	(0.72342)	(0.78641)	(0.00130)
	[716,772]	[624,780]	
College of Agriculture	0.04868	0.03996	0.00872***
	(0.21519)	(0.19586)	(0.00036)
	[716,772]	[624,780]	
College of Business	0.08634	0.13210	-0.04576***
	(0.28087)	(0.33860)	(0.00053)
	[716,772]	[624,780]	
College of Engineering	0.01862	0.08211	-0.06349***
	(0.13518)	(0.27454)	(0.00037)
	[716,772]	[624,780]	
College of Human Services	0.15535	0.09132	0.06402***
	(0.36224)	(0.28807)	(0.00057)
	[716,772]	[624,780]	
Intra-University	0.04376	0.03303	0.01073***
	(0.20456)	(0.17872)	(0.00033)
	[716,772]	[624,780]	
College of Liberal Arts	0.33630	0.30558	0.03073***
	(0.47244)	(0.46065)	(0.00081)
	[716,772]	[624,780]	
College of Natural Resources	0.03534	0.05852	-0.02318***
	(0.18464)	(0.23472)	(0.00036)
	[716,772]	[624,780]	
College of Natural Sciences	0.23166	0.23417	-0.00251***
	(0.42189)	(0.42348)	(0.00073)
	[716,772]	[624,780]	
Collge of Veterinary Sciences	0.04395	0.02321	0.02074***
	(0.20498)	(0.15057)	(0.00031)
	[716,772]	[624,780]	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.
Number of observations in brackets.

Table 2: Fraction of positive and negative grade changes by college and student sex

Grade Change	Upward			Downward		
Students	Female	Male	Difference	Female	Male	Difference
Agriculture	0.00487	0.00296	0.00191***	0.00009	0.00004	0.00005***
	(0.06964)	(0.05437)	(0.00053)	(0.00930)	(0.00634)	(0.00007)
	[34,886]	[24,963]		[34,719]	[24,890]	
Business	0.00404	0.00500	-0.00096***	0.00005	0.00002	0.00002***
	(0.06343)	(0.07056)	(0.00036)	(0.00698)	(0.00494)	(0.00003)
	[61,884]	[82,530]		[61,637]	[82,119]	
Engineering	0.00397	0.00491	-0.00094	0.00015	0.00043	-0.00028
	(0.06290)	(0.06993)	(0.00067)	(0.01227)	(0.02076)	(0.00019)
	[13,345]	[51,280]		[13,294]	[51,050]	
Human Science	0.00503	0.00558	-0.00055	0.00042	0.00125	-0.00084
	(0.07075)	(0.07449)	(0.00037)	(0.02037)	(0.03535)	(0.00014)
	[111,304]	[56,987]		[110,790]	[56,740]	
Liberal Arts	0.00422	0.00538	-0.00117***	0.00015	0.00025	-0.00010***
	(0.06479)	(0.07316)	(0.00021)	(0.01225)	(0.01573)	(0.00004)
	[241,016]	[190,871]		[240,036]	[189,891]	
Natural Resources	0.00426	0.00498	-0.00071	0.00008	0.00005	0.00002
	(0.06516)	(0.07038)	(0.00056)	(0.00890)	(0.00741)	(0.00007)
	[25,328]	[36,560]		[25,222]	[36,380]	
Natural Sciences	0.00325	0.00430	-0.00105***	0.00032	0.00029	0.00003***
	(0.05694)	(0.06544)	(0.00022)	(0.01789)	(0.01698)	(0.00006)
	[165,996]	[146,263]		[165,509]	[145,676]	
Veterinary	0.00337	0.00338	-0.00001	0.00006	0.00014	-0.00007
	(0.05791)	(0.05804)	(0.00058)	(0.00798)	(0.01176)	(0.00009)
	[31,500]	[14,499]		[31,396]	[14,452]	
Intra-University	0.00293	0.00228	0.00066	0.00000	0.00015	-0.00015
	(0.05408)	(0.04767)	(0.00046)	(0.00000)	(0.01207)	(0.00007)
	[31,366]	[20,635]		[31,274]	[20,591]	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.
Number of observations in brackets.

Table 3: Fraction of upward grade changes by gender of students and instructors

	Female Instructor	Male Instructor	Difference
Female Student	0.00412	0.00408	0.000046
	(0.06407)	(0.06371)	(0.000156)
	[335,738]	[335,394]	
Male Student	0.00451	0.00501	-0.000499***
	(0.06703)	(0.07062)	(0.000186)
	[227,750]	[357,090]	
Difference	-0.000391**	-0.000937***	0.000545**
	(0.000177)	(0.000162)	(0.000241)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.
Number of observations in brackets.

Table 4: Fraction of downward grade changes by gender of students and instructors

	Female Instructor	Male Instructor	Difference
Female Student	0.00022	0.00021	0.000018
	(0.01497)	(0.01437)	(0.000036)
	[334,429]	[334,096]	
Male Student	0.00031	0.00032	-0.000011
	(0.01769)	(0.01799)	(0.000048)
	[226,793]	[355,415]	
Difference	-0.000089**	-0.000117***	0.000028
	(0.000044)	(0.000039)	(0.000059)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.
Number of observations in brackets.

Table 5: Grade transition matrix conditional of all grade changes (measured in percent)

	Final Grade									
Initial Grade	A+	A	A-	B+	B	B-	C+	C	D	F
A+	0.00	0.47	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00
A	**0.91**	0.00	0.16	0.13	0.50	0.03	0.00	0.28	0.06	0.06
A-	0.50	**10.12**	0.00	0.00	0.09	0.03	0.00	0.03	0.00	0.00
B+	0.31	2.67	**5.18**	0.00	1.82	0.00	0.16	0.03	0.03	0.00
B	0.22	**12.76**	2.17	**4.59**	0.00	0.09	0.13	0.22	0.06	0.09
B-	0.00	0.31	0.69	0.60	**7.63**	0.00	0.03	0.09	0.00	0.06
C+	0.00	0.16	0.28	0.57	2.20	**3.83**	0.00	0.91	0.00	0.06
C	0.09	1.95	0.31	0.22	**10.49**	1.76	**2.20**	0.00	0.25	0.00
D	0.03	0.44	0.06	0.13	2.04	0.63	0.79	**11.15**	0.00	0.09
F	0.06	0.44	0.00	0.03	0.60	0.06	0.13	1.70	**2.95**	0.00

Table 6: Regression results for upward grade changes

	Dependent Variable: Upward Grade Change $\in \{0,1\}$						
Panel A	[1]	[2]	[3]	[4]	[5]	[6]	[7]
$Male_i$	0.000749***	0.000767***	0.000702***	0.000725***	0.000743***	0.000706***	0.000805***
	(0.000118)	(0.000120)	(0.000123)	(0.000118)	(0.000118)	(0.000117)	(0.000118)
N	1,341,213	1,341,213	1,341,213	1,341,213	1,341,213	1,341,213	1,341,213
Panel B	Omitted reference group: $Female_iFemale_j$						
$Female_iMale_j$	-0.000046	0.000152	0.000018	-0.000095	-0.000063	-0.000428***	-0.000396**
	(0.000155)	(0.000155)	(0.000165)	(0.000158)	(0.000155)	(0.000155)	(0.000155)
$Male_iFemale_j$	0.000391**	0.000484***	0.000495***	0.000372**	0.000398**	0.000180	0.000156
	(0.000181)	(0.000183)	(0.000185)	(0.000181)	(0.000181)	(0.000181)	(0.000182)
$Male_iMale_j$	0.000890***	0.001052***	0.000867***	0.000829***	0.000870***	0.000605***	0.000635***
	(0.000164)	(0.000167)	(0.000181)	(0.000165)	(0.000164)	(0.000163)	(0.000163)
N	1,255,972	1,255,972	1,255,972	1,255,972	1,255,972	1,255,972	1,255,972
F test: Coefficients of all gender pairs = 0	13.671	15.068	11.094	12.738	13.314	13.694	13.443
$(p-value)$	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Control for:							
College		Y					
Department			Y				
Instructor position				Y			
Student class standing					Y		
Gap between GPA and grade						Y	
GPA and grade							Y

$p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the student level. $Male_i$ is an indicator for male student i. $Female_iFemale_j$ is the gender interaction term between female student i and female instructor j, $Female_iMale_j$ is the gender interaction term between female student i and male instructor j, $Male_iFemale_j$ is the gender interaction term between male student i and female instructor j, $Male_iMale_j$ is the gender interaction term between male student i and male instructor j.

Table 7: Marginal effects for upward grade changes

	Dependent Variable: Positive Grade Change $\in \{0,1\}$		
	OLS	Probit	Logit
Panel A			
$Male_i$	0.000790***	0.0008262***	0.0008516***
	(0.000124)	(0.0001278)	(.0001289)
N	1,341,213	1,294,178	1,294,178
Panel B	Omitted reference group: $Female_i Female_j$		
$Female_i Male_j$	-0.000263	-0.0004601***	-0.0004329**
	(0.000166)	(0.0001769)	(0.0001799)
$Male_i Female_j$	0.000440**	0.0003789*	0.0004184**
	(0.000185)	(0.0002029)	(0.0002069)
$Male_i Male_j$	0.000750***	0.0006347***	0.0006936***
	(0.000183)	(0.0001958)	(0.0002008)
N	1,255,972	1,212,186	1,212,186

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the student level. $Male_i$ is an indicator for male student i. $Female_i Female_j$ is the gender interaction term between female student i and female instructor j, $Female_i Male_j$ is the gender interaction term between female student i and male instructor j, $Male_i Female_j$ is the gender interaction term between male student i and female instructor j, $Male_i Male_j$ is the gender interaction term between male student i and male instructor j. All models controls for colleges, departments, instructor positions, student class standings, GPA, and grade.

Table 8: Partial identification for upward grade changes

	Dependent Variable: Positive Grade Change $\in \{0,1\}$		
Regressor	$Female_i Male_j$	$Male_i Female_j$	$Male_i Male_j$
Upper bound	-0.0001116	0.0005583***	0.0008871***
	(0.0001604)	(.0001706)	(.0001768)
Point estimate	-0.000263	0.000440**	0.00075***
	(0.000166)	(0.000185)	(0.000183)
Lower bound	-0.000572***	0.0004221**	0.0004887***
	(0.0001572)	(0.0001842)	(.0001747)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in clustered at the student level. Coefficients from each row comes from a separate regression analysis. Upper-bound estimates assume that all observations with missing value in instructor's sex are from female instructors, and the lower-bound estimates treat all missing values in instructor's sex as male. $Male_i$ is an indicator for male student i. $Female_i Female_j$ is the gender interaction term between female student i and female instructor j, $Female_i Male_j$ is the gender interaction term between female student i and male instructor j, $Male_i Female_j$ is the gender interaction term between male student i and female instructor j, $Male_i Male_j$ is the gender interaction term between male student i and male instructor j. All models controls for colleges, departments, instructor positions, student class standings, GPA, and grade.

Table 9: Sensitivity analysis for upward grade changes

	Dependent Variable: Upward Grade Change $\in \{0,1\}$				
Panel A	Baseline	Excl. students w/2+ regrades [a]	Instructors FE [b]	Exclude A+ and A	Exclude F
$Male_i$	0.000790***	0.000650***	0.000677***	0.001135***	0.000810***
	(0.000124)	(0.000108)	(0.000123)	(0.000184)	(0.000124)
N	1,341,213	1,327,826	1,341,213	876,959	1,293,734
Panel B	Omitted reference group: $Female_iFemale_j$				
$Female_iMale_j$	-0.000263	-0.000266*		-0.000407	-0.000331**
	(0.000166)	(0.000156)		(0.000267)	(0.000167)
$Male_iFemale_j$	0.000440**	0.000281*	0.000464**	0.000587**	0.000419**
	(0.000185)	(0.000166)	(0.000186)	(0.000275)	(0.000187)
$Male_iMale_j$	0.000750***	0.000643***	0.000843***	0.001110***	0.000743***
	(0.000183)	(0.000167)	(0.000171)	(0.000278)	(0.000185)
N	1,255,972	1,243,458	1,255,972	819,416	1,211,694
F test: Coefficients of all gender pairs = 0	13.317	12.858	14.898	12.980	14.668
p-value	0.00000	0.00000	0.00000	0.00000	0.00000

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the student level. $Male_i$ is an indicator for male student i. $Female_iFemale_j$ is the gender interaction term between female student i and female instructor j, $Female_iMale_j$ is the gender interaction term between female student i and male instructor j, $Male_iFemale_j$ is the gender interaction term between male student i and female instructor j, $Male_iMale_j$ is the gender interaction term between male student i and male instructor j. All models controls for colleges, departments, instructor positions, student class standings, GPA, and grade.

a. Exclude students who received multiple upward grade changes.

b. Control for instructors fixed effects. The omitted reference group is female students, regardless of the sex of instructors.

Table 10: Regression results for downward grade changes

	Dependent Variable: Downward Grade Change $\in \{0,1\}$						
Panel A	[1]	[2]	[3]	[4]	[5]	[6]	[7]
$Male_i$	0.000103***	0.000135***	0.000024	0.000103***	0.000100***	0.000104***	0.000090***
	(0.000029)	(0.000030)	(0.000027)	(0.000029)	(0.000029)	(0.000029)	(0.000029)
N	1,335,666	1,335,666	1,335,666	1,335,666	1,335,666	1,335,666	1,335,666
Panel B	Omitted reference group: $Female_iFemale_j$						
$Female_iMale_j$	-0.000018	0.000076**	0.000031	-0.000005	-0.000024	-0.000004	0.000005
	(0.000036)	(0.000036)	(0.000038)	(0.000036)	(0.000036)	(0.000036)	(0.000036)
$Male_iFemale_j$	0.000089**	0.000162***	0.000045	0.000085*	0.000089*	0.000096**	0.000082*
	(0.000045)	(0.000048)	(0.000045)	(0.000045)	(0.000045)	(0.000045)	(0.000045)
$Male_iMale_j$	0.000099**	0.000192***	0.000038	0.000112***	0.000088**	0.000109***	0.000105***
	(0.000040)	(0.000042)	(0.000042)	(0.000041)	(0.000039)	(0.000040)	(0.000039)
N	1,250,733	1,250,733	1,250,733	1,250,733	1,250,733	1,250,733	1,250,733
F test: Coefficients of all gender pairs = 0	4.118	7.556	0.451	4.095	3.911	4.194	3.382
p-value	0.00627	0.00005	0.71636	0.00647	0.00836	0.00564	0.01737
Control for:							
College		Y					
Department			Y				
Instructor position				Y			
Student class standing					Y		
Gap between GPA and grade						Y	
GPA and grade							Y

$p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the student level. $Male_i$ is an indicator for male student i. $Female_iFemale_j$ is the gender interaction term between female student i and female instructor j, $Female_iMale_j$ is the gender interaction term between female student i and male instructor j, $Male_iFemale_j$ is the gender interaction term between male student i and female instructor j, $Male_iMale_j$ is the gender interaction term between male student i and male instructor j.

Table 11: Sensitivity analysis for downward grade changes

	Dependent Variable: Downward Grade Change $\in \{0,1\}$				
Panel A	Baseline	Excl. students w/2+ regrades[a]	Instructors FE[b]	Exclude A+ and A	Exclude F
$Male_i$	-0.000002	-0.000012	-0.000011	0.000044	-0.000004
	(0.000027)	(0.000027)	(0.000026)	(0.000032)	(0.000028)
N	1,335,666	1,335,342	1,335,666	871,363	1,288,530
Panel B	Omitted reference group: $Female_iFemale_j$				
$Female_iMale_j$	0.000037	0.000038		0.000125***	0.000040
	(0.000039)	(0.000039)		(0.000046)	(0.000040)
$Male_iFemale_j$	0.000026	0.000006	-0.000044	0.000135***	0.000024
	(0.000044)	(0.000044)	(0.000043)	(0.000050)	(0.000046)
$Male_iMale_j$	0.000016	0.000013	0.000019	0.000098**	0.000015
	(0.000042)	(0.000041)	(0.000036)	(0.000048)	(0.000043)
N	1,250,733	1,250,421	1,250,733	814,112	1,206,771
F test: Coefficients of all gender pairs = 0	0.333	0.376	0.659	3.783	0.357
p-value	0.80117	0.77017	0.51724	0.00998	0.78401

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the student level. $Male_i$ is an indicator for male student i. $Female_iFemale_j$ is the gender interaction term between female student i and female instructor j, $Female_iMale_j$ is the gender interaction term between female student i and male instructor j, $Male_iFemale_j$ is the gender interaction term between male student i and female instructor j, $Male_iMale_j$ is the gender interaction term between male student i and male instructor j. All models controls for colleges, departments, instructor positions, student class standings, GPA, and grade.

a. Exclude students who received multiple upward grade changes.

b. Exclude instructors who changed more than 10 percent of grade records upward. The omitted reference group is female students, regardless of the sex of instructors.

Table 12: Grade changes by initial grade, all students

	No regrades	Regrades		
Initial Grade	Grade	Inital Grade	Final Grade	Grade Change
A+,A,A-	3.94300	3.73442	3.88080	0.14638***
	(0.12543)	(0.13388)	(0.41812)	(0.01407)
	[560,052]	[973]	[973]	
B+,B,B-	3.02165	3.01553	3.61094	0.59541***
	(0.18587)	(0.22793)	(0.46662)	(0.01033)
	[461,300]	[2,529]	[2,529]	
C+,C	2.06565	2.09580	2.96151	0.86571***
	(0.13273)	(0.15111)	(0.49148)	(0.01316)
	[210,722]	[1,527]	[1,527]	
D	1.00000	1.00000	2.27555	1.27555***
	(0.00000)	(0.00000)	(0.55130)	(0.01888)
	[56,117]	[853]	[853]	
F	0.00000	0.00000	1.95631	1.95631***
	(0.00000)	(0.00000)	(1.03438)	(0.05585)
	[47,136]	[343]	[343]	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.
Number of observations in brackets.

Table 13: Grade changes by initial grade, female students

	No regrades	Regrades		
Initial Grade	Grade	Inital Grade	Final Grade	Grade Change
A+,A,A-	3.94443	3.73592	3.89929	0.16337***
	(0.12416)	(0.13503)	(0.33819)	(0.01558)
	[332,938]	[546]	[546]	
B+,B,B-	3.02828	3.02437	3.66895	0.64458***
	(0.18562)	(0.22160)	(0.40774)	(0.01309)
	[235,878]	[1,257]	[1,257]	
C+,C	2.06479	2.08466	3.00984	0.92518***
	(0.13207)	(0.14539)	(0.47075)	(0.01834)
	[98,446]	[722]	[722]	
D	1.00000	1.00000	2.31782	1.31782***
	(0.00000)	(0.00000)	(0.58819)	(0.03083)
	[25,833]	[364]	[364]	
F	0.00000	0.00000	2.15692	2.15692***
	(0.00000)	(0.00000)	(1.09633)	(0.08863)
	[20,635]	[153]	[153]	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Number of observations in brackets.

Table 14: Grade changes by initial grade, male students

	No regrades	Regrades		
Initial Grade	Grade	Inital Grade	Final Grade	Grade Change
A+,A,A-	3.94089	3.73251	3.85715	0.12465***
	(0.12724)	(0.13253)	(0.50157)	(0.02511)
	[227,114]	[427]	[427]	
B+,B,B-	3.01472	3.00680	3.55362	0.54682***
	(0.18589)	(0.23378)	(0.51202)	(0.01578)
	[225,422]	[1,272]	[1,272]	
C+,C	2.06640	2.10580	2.91817	0.81237***
	(0.13330)	(0.15548)	(0.50574)	(0.01865)
	[112,276]	[805]	[805]	
D	1.00000	1.00000	2.24409	1.24409***
	(0.00000)	(0.00000)	(0.52054)	(0.02354)
	[30,284]	[489]	[489]	
F	0.00000	0.00000	1.79476	1.79476***
	(0.00000)	(0.00000)	(0.95428)	(0.06923)
	[26,501]	[190]	[190]	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.
Number of observations in brackets.