The effects of Greek affiliation on academic performance*

Andrew De Donato[†] and James Thomas[‡]

July 8, 2015

Abstract

This paper uses administrative data from Duke University and a difference-in-differences approach to estimate the effects of Greek affiliation on academic performance. We find evidence of strong negative effects in some periods but smaller effects in others: fraternity affiliation hurts performance by .32 standard deviations in the Freshman Spring semester when students complete new member education; sorority affiliation hurts performance by .22 standard deviations in Spring semesters after Freshman year when students recruit and educate new members. We estimate both a 'partial effect' which controls for changes in course choice behavior and a 'total effect' which allows Greek affiliation to influence course choice behavior. The difference between these suggests students mitigate the negative effects of affiliation by selecting into courses which grade more leniently. We account for substantial censoring of grades at 4.0 and show that ignoring censoring leads to under-estimates of the negative effects of affiliation. We also document substantial heterogeneity in treatment effects by initial preparation of students and social status of Greek organizations.

*We are especially grateful to Peter Arcidiacono and Michelle Connolly for ongoing advice on this project and to Ken Spenner for providing access to the Campus Life and Learning data. We are also grateful to Kehinde Ajayi, Connel Fullenkamp, Robert Garlick, Marie Hull, Margaux Luflade, Arnaud Maurel, Kate Vyborny, Maria Zhu, and participants in the Duke University Labor Lunch Seminar for helpful comments. All remaining errors are our own.

[†]Duke University; andrew.de.donato@duke.edu.

[‡]Duke University; james.r.thomas@duke.edu.

1 Introduction

Universities across the United States are at a crossroads: what should be done about fraternities and sororities? In November, 2014, West Virginia University suspended all fraternity and sorority activities after an 18 year old student died at a fraternity house (Johnson, 2014). The University of Virginia suspended all fraternity activities for six weeks and overhauled its policies related to Greek organizations following a Rolling Stones article depicting sexual assault which was later retracted (DeBonis and Shapiro, 2014). Dartmouth College faculty recently voted to eliminate the college's Greek system after a former fraternity member published a book with disturbing revelations about Greek life at Dartmouth (Jacobs, 2014). Across the country, similar incidents have occurred and similar policy decisions have followed (Perez-Pena and Yaccino, 2014).

Those opposed to Greek organizations portray them as hotbeds of drinking, drug use, sexual assault, hazing, racism, and elitism. This view holds that Greek organizations are the antithesis of what higher education should strive towards. Those in support of Greek organizations note their participation in community service and charity fund raising, their positive contributions to college social life, and their outsized role in alumni giving. This view claims that, while a few atrocities grab headlines, Greek organizations are nonetheless valuable institutions on campuses. There clearly exists a vibrant debate about whether Greek organizations detract from or contribute to the missions of universities. This paper sheds light on the effect of Greek organizations on a principal component of university missions: academic development of students.

Specifically, this paper quantifies the effects of joining a Greek organization on academic performance. We find evidence of large negative effects in certain semesters but small and insignificant effects in other semesters. Male students experience large negative effects in the Freshman Spring semester (when recruitment and new member education occurs) and modest negative effects in subsequent semesters.[1] Comparatively, female students experience modest negative effects in

[1] In a preferred specification, the negative effect in Freshman Spring is .216 grade points (31.7% of a standard deviation), the negative effect in Fall semesters after Freshman year is .132 grade points (19.4% of a standard deviation), and the negative effect in Spring semesters after Freshman year is .089 grade points (13.1% of a standard deviation). Grade points use the following scale: 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).

Spring semesters after Freshman year but small and insignificant effects in other semesters.[2] This suggests Greek recruitment and initiation is a significant distraction from schoolwork—for men, the burden falls primarily on Freshmen being initiated; for women, the burden is on upper class students responsible for initiation.

We also find evidence that effects vary across students and depend on characteristics of the Greek organization. First, we find students with SAT scores below the median experience strong negative effects while students above the median are less affected. This is especially true for male students. Second, we find Greek organizations with higher social status have strong negative effects during new member education but smaller effects in subsequent years. Comparatively, organizations with lower social status have smaller effects during new member education but larger effects in later years.

These findings contribute to a growing literature on the determinants of academic outcomes in higher education by showing Greek affiliation can have sizable negative effects on performance.[3] In addition, the results offer insights into ways administrators can reform Greek policies to reduce their impacts. First, Greek recruitment and new member education appears to be taking a significant toll on academic performance. For men, the burden falls on new members being educated; for women, the burden falls on upper class members who recruit and educate new members. University administrators may consider reforming recruitment and new member education to make it less distracting or moving it so it does not conflict with coursework. Second, less prepared students are being hurt by Greek organizations more than other students. Universities could monitor the academic performance of 'at risk students' in each organization and reward (or punish) the entire organization based on the performance of these students. More drastically, they could set eligibility cutoffs to prevent at risk students from ever affiliating.

Our analysis focuses on Duke University where an unusual policy of 'deferred affiliation' al-

[2]In a preferred specification, the negative effect in Spring semesters after Freshman year is .146 grade points (21.5% of a standard deviation) but only .082 grade points (12.1% of a standard deviation) in Freshman Spring and .095 grade points (14.0% of a standard deviation) in Fall semesters after Freshman year.

[3]Notable works include but are not limited to: Bound, Lovenheim, and Turner (2010); Bound and Turner (2007); Light and Strayer (2000); Stinebrickner and Stinebrickner (2004 and 2003).

lows us to employ a difference-in-differences (DinD) approach to identify the effects of Greek affiliation on academic performance. Deferred affiliation prevents students from joining Greek organizations in the Fall semester of Freshman year. This allows us to observe one semester of baseline grades for all students before selection into Greek organizations occurs. Intuitively, we compare the change in academic performance for Greek students before and after they affiliate to the change in performance over the same period for independent students. This double difference approach controls for the possibility that students who join Greek organizations may be initially more or less prepared for academic success than students who remain independent. This is an important consideration especially for female students: even with robust controls for demographic characteristics and observed measures of initial preparation, future sorority members outperform female students who remain independent by .122 grade points in the Fall semester of Freshman year.

Our analysis also controls for changes in course choice behavior which result from Greek affiliation. Existing literature shows grading standards differ across courses making it possible to selectively choose leniently or harshly graded courses (Johnson, 2003; Sabot and Wakeman-Linn, 1999). If joining a Greek organization leads students to selectively choose easier courses this will mitigate the negative effects of Greek affiliation on grades but only in a superficial manner. We develop methodologies to estimate both a 'total effect' which allows for changes in course choice behavior and a 'partial effect' which controls for the grading leniency of chosen courses. Furthermore, we show the difference between these effects can be directly interpreted as the effect of Greek affiliation on the grading leniency of chosen courses. Our estimates suggest these behavioral responses have statistically significant mitigating effects in certain settings; ignoring these effects leads to under-estimates of the effects of Greek affiliation on academic performance. Our analysis of the total effect, partial effect, and leniency effect provides a more complete picture of the effects of Greek affiliation on academic performance: the partial effect measures pure change in academic performance, the total effect measures the change in expected outcomes, and the leniency effect measures how Greek affiliation affects grading leniency of chosen courses.

Our methodology also adjusts for the contaminating effects of censoring at 4.0 grade points. In our sample, 33.4% of assigned grades are worth 4.0 grade points. This represents substantial right censoring which causes well known econometric concerns (Tobin, 1958). We extend traditional DinD estimation techniques to accommodate censoring of the dependent variable. We compare our estimates to OLS estimates and find ignoring censoring leads to treatment effect estimates which are attenuated by as much as 22.9%. Pervasive grade inflation causes substantial right censoring of grades at many post-secondary institutions. Our findings suggest ignoring this censoring may lead to large attenuation bias in certain settings.

There exists a rich literature examining the effects of Greek affiliation on many aspects of university life. Some studies find Greek organizations have negative effects such as encouraging binge drinking (Wechsler, Kuh, and Davenport, 1996), gambling (Rockey et al., 2005), drug use (Aikins, 2011), and cheating (McCabe and Trevino, 1997; Passow et al., 2006; Williams and Janosik, 2007). Other studies find Greek organizations have positive effects such as increasing self-efficacy (Hayek et al., 2002; Thompson, Oberle, and Liley, 2011), developing interpersonal skills (Pike, 2003; 2004; Pike and Askew, 1990), and increasing civic engagement (Astin, 1993; Pascarella and Terenzini, 2005).

The two studies most relevant to our work are Grubb (2006) and Walker, Martin, and Hussey (2014) which examine the effects of Greek affiliation on academic outcomes. Grubb (2006) finds fraternity involvement decreases final GPA by 2.2% but finds no effect of sorority involvement using data from the University of Delaware. Grubb (2006) controls for SAT scores and state of residence but does not control for other confounding factors. Walker, Martin, and Hussey (2014) use data from Duke University—the same data used in our study—and a propensity score matching estimator to estimate the effects of Greek affiliation on social networks, time allocation, social identities, satisfaction with college, and academic performance. The authors find interesting effects on many other outcomes but find insignificant effects of Greek affiliation on academic performance. The selection model used in Walker, Martin, and Hussey (2014) employs detailed baseline controls including demographic variables, measures of academic preparation, and responses to survey questions regarding identity and expectations; however, even with these rich controls their model achieves a pseudo R^2 of only .11 suggesting there are many unobserved factors affecting selection which may confound estimates of treatment effects.

Both Grubb (2006) and Walker, Martin, and Hussey (2014) rely on selection-on-observables approaches for estimating treatment effects. Our DinD approach adjusts for both selection on observed factors and selection on unobserved factors. We find evidence that selection on unobserved factors can have large confounding effects making this adjustment crucial for providing accurate results. Additionally, we are the first to account for censoring of grades at 4.0 and for changes in course choice behavior which result from Greek affiliation. Once again, we show both of these adjustments lead to substantially different estimates of treatment effects. Finally, we extend Grubb (2006) and Walker, Martin, and Hussey (2014) by estimating separate treatment effects by gender, academic semesters, initial preparation of students, and social status of the Greek organizations. We find evidence of interesting heterogeneity in effects providing a more complete analysis of the effects of Greek affiliation on academic outcomes.

The remainder of this paper is organized as follows: Section 2 provides background information about Greek life at Duke and Duke's deferred affiliation policy which is crucial for our identification strategy. Section 3 introduces our administrative data and discusses the strengths of these data for our study. Section 4 analyses selection into Greek organizations with descriptive statistics for Greek and non-Greek populations. Section 5 describes our methods for controlling for this non-random selection and for accounting for changes in course choice behavior and censoring at 4.0. Section 6 presents the details of our empirical specification and discusses threats to identification. Section 7 presents results on the effects of Greek affiliation on academic outcomes. Section 8 concludes.

2 Institutional Background: Greek life at Duke University

Greek life is an important component of campus culture at Duke University. There are 34 Greek organizations including 15 historically white fraternities, 10 historically white sororities, 5 African American fraternities and sororities, and 4 multicultural fraternities and sororities. 14.9% of all undergraduates ever affiliate with a traditional fraternity, 17.4% of all undergraduates ever affiliate with a traditional sorority, 1.9% of all undergraduates ever affiliate with an African American fraternity or sorority, and 0.7% of all undergraduates ever affiliate with a multicultural fraternity or sorority. These levels of involvement are quite high by national standards: Duke University ranks 32 in sorority involvement and 62 in fraternity involvement ("Most Students in Sororities," 2013; "Most Students in Fraternities," 2013).

Administrators at Duke University view Greek organizations as positive institutions which develop individuals and enrich campus culture. Greek organizations are believed to "educate students in the areas of leadership, cultural awareness, personal and group development, scholarship and civic responsibility." Additionally, Greek organizations contribute to campus life by "sponsoring inter-campus events, hosting parties, and performing community service" (Fraternity & Sorority Life, 2015).

However, university administrators are also aware of the concerns surrounding Greek organizations. To give students more time to make informed decisions about affiliating, Duke has a policy of 'deferred affiliation' whereby students cannot affiliate with Greek organizations until the Spring semester of their Freshman year. This is an uncommon policy which provides useful variation for identifying the effects of Greek affiliation on academic outcomes.[4] Intuitively, deferred affiliation allows the econometrician to measure the change in the academic performance of Greek students before and after they affiliate. A difference-in-differences approach compares this change to the change over the same period for independent students to measure the effects of affiliation.[5]

[4] Proponents of deferred affiliation policies argue they give students an opportunity to transition to college life, reduce underage drinking, and give students a chance to make more informed decisions about affiliation. Opponents argue these policies delay the social benefits of Greek affiliation and oppose freedom of association. A policy analysis by Auburn University finds only 4 of 22 universities studied have deferred affiliation policies (Kittle, 2011).

[5] While it is theoretically possible to join Greek organizations after the Spring semester of Freshman year this is

The process of Greek affiliation involves a short period of recruitment followed by a longer period of new member education. During recruitment, prospective students attend events hosted by Greek organizations and an informal process of mutual selection determines which students earn 'bids' to affiliate. In Section 5, we discuss how this selection process complicates identification of the effects of Greek affiliation on academic outcomes and how our difference-in-differences approach handles this complication.

Following recruitment, prospective members complete new member education—a process known as 'pledging'—before they are accepted as full members of the organization. While Duke University has made concerted efforts to make pledging a positive experience, an intransigent culture of hazing still infiltrates new member education (Tuchler, 2012). New member education can be quite challenging for prospective members and can also be taxing on senior members who participate in education events. For this reason, we estimate separate treatment effects for the Freshman Spring semester (when Greek students complete new member education), Fall semesters after Freshman year (when limited recruitment occurs), and Spring semesters after Freshman year (when Greek students may administer new member education). This provides a complete picture of the effects of Greek affiliation over the course of college.

As is the case at many universities, Greek organizations at Duke range from very selective groups which are highly regarded on campus to less selective groups. This social hierarchy is generally agreed upon and known throughout the student body. Terms such as "key three"—referring to the top tier sororities—are used around campus without additional explanation. To assess whether the effects of Greek affiliation on academic outcomes depend on the social status of an organization we categorize organizations into high status and low status based on a perusal of comments in the forum Greekrank.com.[6] High status organizations are those which are commonly

relatively rare (Walker, Martin, and Hussey, 2015).

[6]Greekrank.com is an online forum in which users debate the social hierarchy of Greek organizations at various universities. High status fraternities are: KA Order, Sigma Nu, Delta Sigma Phi, PiKapp, Pi Kappa Alpha, Alpha Tau Omega, Sigma Phi Epsilon, and Delta Tau Delta. High status sororities are: Pi Beta Phi, Tridelt, Kappa Kappa Gamma, Kappa Alpha Theta, Delta Gamma, and Alpha Phi. All ranked organizations are traditional fraternities and sororities. With this categorization, 61% of sorority members belong to high status sororities and 56% of fraternity members belong to high status fraternities.

included in rankings by commenters. Our perusal suggests physical attractiveness of members and quality of social events are important determinants of where an organization sits in the social hierarchy. While this measure is purely qualitative, we note that the debates on Greekrank.com are usually over minor differences in rank. As such, we feel confident that our categorization is consistent with conventional campus wisdom. Furthermore, this measure is only used to estimate heterogeneous effects by organization social status and does not influence our main results.

3 Duke University Data

To estimate the effects of Greek affiliation on academic outcomes, we employ administrative and survey data from Duke University which includes full academic transcripts, administrative and data on Greek affiliation, information used to make admissions decisions, and self-reported demographic information. These data are available for students who participated in the Campus Life and Learning (CLL) Survey which followed sub-samples of the 2001 and 2002 entering undergraduate cohorts at Duke University.[7]

These data are well suited for our analysis for two important reasons: First, full academic transcripts allow us to observe individual grades and identify the courses they were earned in. As we explain in Section 5, this allows us to control for different grading standards across courses and to adjust for the fact that grades are capped at 4.0. Previous literature uses aggregate measures of academic performance and thus does not control for grading standards or account for censoring (Grubb, 2006; Walker, Martin, and Hussey, 2015). In Section 7, we show both of these omissions lead to under-stating the negative effects of Greek affiliation on academic outcomes.

Second, full academic transcripts allow us to construct data panels of the same student's grades over their entire college career. This allows us to employ a difference-in-differences approach to control for unobserved factors which may affect both academic outcomes and decisions to affiliate with Greek organizations. Additionally, it allows us to measure the effects of Greek affiliation

[7]For detailed reports on the CLL data see Bryant, Spenner, and Martin (2006 and 2007).

at various points throughout a student's college career. In Section 7, we show the importance of controlling for unobserved factors and illustrate interesting dynamics in the effects of Greek affiliation over the course of college.

The CLL Survey originally contacted 1,536 students in the 2001 and 2002 entering undergraduate cohorts at Duke University. Of these students, 1,132 ever gave consent to have their confidential records used for research purposes. Arcidiacono, Aucejo, and Spenner (2012) show non-consenters have lower SAT scores, have better educated parents, are more likely to be from private schools, and have slightly lower grade point averages than respondents. However, they note the differences are quite small and thus conclude the CLL sample is generally representative of the 2001 and 2002 cohorts.[8]

To ensure the composition of the Greek and non-Greek samples are consistent over time we further restrict the sample to students who appear in every academic semester except possibly Junior Fall.[9] While dropout is rare at Duke University, 124 of the 1,132 CLL consenters are missing from at least one semester other than Junior Fall. Excluding these 124 students results in a final sample of 1,008 unique students and 29,001 student-course observations. In an online appendix, we show that our estimation sample is indistinguishable from the full sample.[10]

To define our measure of Greek affiliation we use administrative records which report whether a student was affiliated at the end of his or her Freshman year. While it is theoretically possible to join Greek organizations after the Spring semester of Freshman year this is relatively rare (Walker, Martin, and Hussey, 2015). Categorizing these unobserved late joiners as independent introduces attenuation bias in estimates of the effects of Greek affiliation in Sophomore year and beyond. It is also possible for members to disassociate with their Greek organizations. Data for assessing the frequency of disassociation are not available; however, anecdotal evidence suggests this is rare. In

[8]The CLL intentionally over-sampled minority students; as such, survey weights are required to make the CLL sample representative of the 2001 and 2002 cohorts. The survey weights we employ also adjust for non-response to improve the representativeness of the sample.

[9]Junior Fall is excluded from our analysis because a large number of students study abroad without earning transferable grades in this semester.

[10]Importantly, the fraction of men who ever join a fraternity and women who ever join a sorority is identical in the full and restricted samples. In the full sample, 32.1% of men ever join a fraternity and 37.4% of women ever join a sorority; in the restricted sample, the respective proportions are 31.9% and 37.8%.

the spirit of "intention to treat analysis" (Angrist, Imbens, and Rubin, 1996), our estimates measure the effects of ever affiliating with a Greek organization. Disassociations attenuate our intention to treat estimates relative to the effects of being currently affiliated with a Greek organization.

4 Descriptive Statistics

Table 1 describes the Greek and independent populations at Duke. In general, Greek students are whiter and wealthier than independents; however, Greeks and independents have very similar admissions scores.

The racial gap is particularly pronounced for female students: sorority members are 24.5 percentage points more likely to be white than female independents while fraternity members are only 11.4 percentage points more likely to be white than male independents. Sorority members are also more likely to be legacy students or from a private high school than female independents.[11] For male students, these differences are small and insignificant.

While Duke students generally come from wealthy and highly educated families, Greek students at Duke come from especially wealthy and educated families. 42.2% of sorority members and 44.5% of fraternity members have at least one parent with an advanced professional degree compared to 34% of female independents and 37.1% of male independents.[12] Furthermore, 34.5% of sorority members and 39.9% of fraternity members come from families earning over $200,000 per year compared to 21.7% of female independents and 23.8% of male independents.

These large differences in socioeconomic status raise concerns about the common trends assumption of difference-in-differences estimators. As we discuss in Section 6, we address this concern by allowing demographic controls to have differential effects over time. Furthermore, as we show in Section 7, these time varying demographic controls have negligible effects on treatment effect estimates which supports the validity of the common trends assumptions.

Surprisingly, these substantial demographic differences do not correspond to differences in ap-

[11] Legacy students have a close relative who also attended Duke University.

[12] Advanced professional degrees include law degrees, medical degrees, and doctoral degrees.

plication quality. Average SAT scores and internal application assessments are very similar across Greek and independent populations. This is surprising because there is generally a strong positive correlation between family income and SAT scores (Zwick, 2004). As such, Greek students appear to be positively selected on socioeconomic status and negatively selected on academic preparation conditional on socioeconomic status.

Table 1 also includes measures of performance and grading leniency during Freshman year. Performance is measured by grade point average while grading leniency is measured by the average grade of other students in the classes taken by a student.[13] The performance figures show future Greek members outperform students who remain independent before Greek affiliation occurs; however, Greek members lose some (or all) of their advantage after affiliation occurs. In the Fall semester of Freshman year, future sorority members perform significantly better than female students who remain independent. Point estimates suggest future fraternity members also outperform male students who remain independent but the difference is not statistically significant. This suggests Greek members are stronger students than independents on average. In the Spring semester of Freshman year, sorority members continue to outperform female independents but the gap narrows slightly. Conversely, point estimates suggest fraternity members now perform worse than male independents although the difference is still insignificant.

Furthermore, the grading leniency figures suggest Greek students take more difficult courses before affiliation occurs and switch to more lenient courses after affiliating. This suggests changes in raw GPA understate the true effects of affiliation on academic performance. These changes in

[13] More precisely, let Y_{itj} denote the grade earned by student i in course j and semester t and let d_{itj} indicate whether student i took course j in semester t.

Performance of student i in semester t is given by:

$$\bar{Y}_{it} = \frac{\sum_j d_{itj} Y_{itj}}{\sum_j d_{itj}}$$

Grading leniency of student i in semester t is given by:

$$\gamma_{it} = \frac{\sum_j d_{itj} \left(\frac{\left\{ \sum_{i'=1}^{N} d_{i'tj} Y_{i'tj} \right\} - Y_{itj}}{\left\{ \sum_{i'=1}^{N} d_{i'tj} \right\} - 1} \right)}{\sum_j d_{itj}}$$

Sampling weights are used to adjust both of these figures for stratified sampling and non-response in the CLL.

course choice behavior are especially pronounced for fraternity members: In the Fall semester of Freshman year, future fraternity members take courses which are .025 grade points more difficult than courses taken by male independents; however, in the Spring semester of Freshman year, fraternity members choose courses which are .028 grade points more lenient than courses taken by male independents. In both cases, the differences are at the margin of statistical significance. The figures suggest sorority members also switch from more challenging to less challenging courses after affiliating but the differences are smaller and statistically insignificant.

These performance and grading leniency figures demonstrate the intuition of our difference-in-differences approach and show the limitations of studies which rely on SAT scores and other admissions criteria to control for selection. If researchers directly compare Greek and independent GPA after affiliation occurs they will wrongly conclude sorority affiliation helps female students and fraternity affiliation has no effect on male students. This is incorrect because these Greeks were stronger students to begin with. It is also incorrect because it ignores the fact that the Greek students were taking more lenient courses after affiliation. Furthermore, because Greek students do not have advantages in SAT scores or other admissions criteria, controlling for these variables will not account for positive selection into Greek organizations or differences in course choice behavior.

However, comparing the change in average performance before and after affiliation for Greek students to the change in average performance for independent students over the same period suggests that sorority affiliation decreases performance by .05 grade points and fraternity affiliation decreases affiliation by .12 grade points. Additionally, comparing the change in grading leniency before and after affiliation for Greek students to the same change for independents suggests sorority affiliation increases grading leniency by .02 grade points and fraternity affiliation increases grading leniency by .05. This implies Greek affiliation leads to decreases in performance despite increases in grading leniency. Combining these two effects suggests sorority affiliation hurts true performance by .07 grade points and fraternity affiliation hurts true performance by .17 grade points.

There are several limitations of this simple adjusted difference-in-differences: First, the measure of grading leniency is crude because it does not account for differences in student ability. Second, the approach ignores the confounding effects of censoring at 4.0 grade points. Furthermore, it ignores the possibility that large demographic differences between Greeks and independents may affect performance differentially over time. In the following two sections, we develop more robust difference-in-difference estimators which address these (and other) limitations. However, the intuition of comparing Greek performance before and after affiliation to independent performance over the same period to control for selection on unobserved factors and using the performance of other students to adjust for grading leniency continues to apply to our estimators.

5 Identification and Estimation Theory

In this section, we describe our approach for identifying the effects of Greek affiliation on academic performance. We begin by describing how non-random selection confounds identification of these effects and the intuition underlying the Difference-in-Differences (DinD) approach we use to address these selection concerns.

Next we describe identification of two treatment parameters of interest: a 'partial treatment effect' and a 'total treatment effect'. In a setting in which treatment status may lead to behavioral responses, researchers may be interested in both the partial derivative—measuring how treatment status directly affects outcomes holding behavioral responses fixed—and the total derivative—measuring how treatment status and behavioral responses to treatment combine to affect outcomes (Duflo, Glennerster, and Kremer, 2008). This distinction has also been conceptualized as a *ceteris paribus* effect which holds behavioral responses fixed and a *non-ceteris-paribus* effect which includes indirect effects resulting from behavioral responses to treatment (Todd and Wolpin, 2003).

Our partial treatment effect—or *ceteris paribus* effect—controls for changes in course choice behavior possibly resulting from Greek affiliation while our total treatment effect—or *non-ceteris-paribus* effect—measures the combined effect of Greek affiliation and changes in course choice behavior resulting from affiliation. Additionally, we show the difference between these two treatment effects identifies the effect of Greek affiliation on grading leniency.

Following this, we describe how our estimation approach adjusts for censoring of grades at 4.0. When the dependent variable is censored, OLS estimates suffer from attenuation bias. We incorporate censored regression methods into our DinD framework to provide estimates which account for non-random selection into Greek organizations and censoring of grades at 4.0.

For ease of exposition, this section considers a simplified setting with two populations (treated and non-treated) and two periods (pre-treatment and post-treatment). In Section 6, we present our empirical specification which allows treatment effects to vary by academic semester and by gender, initial preparation, and social status of Greek organization.

5.1 Selection into Greek Organizations

As discussed in Section 2, students who elect to participate in Greek recruitment are sorted into organizations through an informal process of mutual selection. Students decide which organizations to express interest in and organizations decide which students to recruit. Both observed and unobserved characteristics likely affect whether a student participates in recruitment, which organizations the student expresses interest in, and which organizations recruit this student. In Section 4, we showed that Greek students are whiter and wealthier than independents; we should also expect there to be differences in unobserved preferences and traits across these populations.

These unobserved preferences and traits may also affect academic outcomes. For example, preferences for partying may make students more likely to join a Greek organization and also less studious. Conversely, high self-esteem may make students more likely to be accepted into a Greek organization and also more likely to succeed in the classroom. Ignoring this selection on unobserved variables may lead to spurious conclusions regarding the effects of Greek affiliation on academic performance.

The unusual institutional environment at Duke University allows us to address selection on unobserved variables using a DinD estimator (Ashenfelter and Card, 1985). Because Greek affiliation at Duke cannot occur until the Spring semester of Freshman year, we observe all students earning one semester of grades before ever joining a Greek organization. Intuitively, this allows us to compare the change in academic performance before and after affiliation for students who ever join a Greek organization to the change in academic performance over the same period for students who remain independent. This double difference accounts for both unobserved differences between Greeks and independents and for general trends in performance across academic semesters.

The critical assumption for this DinD estimator is that students who join a Greek organization would have experienced the same change in performance over time as independents in a hypothetical scenario when they never join Greek organizations. We discuss this crucial assumption in Section 6.

5.2 Partial Treatment Effect

In this subsection, we present a DinD approach to identify the partial treatment effect which controls for changes in course choice behavior resulting from joining a Greek organization. Existing literature shows grading standards differ across courses making it possible to selectively choose leniently or harshly graded courses (Johnson, 2003; Sabot and Wakeman-Linn, 1999). If joining a Greek organization leads students to selectively choose easier courses this will mitigate the negative effects of Greek affiliation on grades but only in a superficial manner. Our partial treatment effect adjusts for this by measuring the effect of Greek affiliation on academic outcomes conditional on grading leniency. As such, it measures the change in grade point average (GPA) that a prospective Greek student should expect to experience if she joins a Greek organization but continues choosing courses like an independent student.[14]

Let Y_{itj} represent the grade earned by student i in period t in course j. For simplicity, periods are divided into pre-treatment ($t = 0$) and post-treatment ($t = 1$). Let D_{it} indicate whether individual i is treated in period t so $D_{i0} = 0$ for all students, $D_{i1} = 0$ for students who are not treated, and $D_{i1} = 1$

[14]Students may also respond to Greek affiliation by registering for fewer courses per term. Our partial effect estimator can be adjusted to also control for these responses by including controls for semester course load. Preliminary results suggested this is not an important mechanism so we have excluded it for clarity.

for students who are treated. Following Ashenfelter and Card (1985), suppose earned grades can be additively separated into a period specific effect δ_t, the effect of treatment α, observed factors X_{it}, an individual effect η_i, a measure of course specific grading leniency γ_j, and idiosyncratic noise υ_{itj}:

$$Y_{itj} = \delta_t + \alpha D_{it} + X_{it}\beta + \eta_i + \gamma_j + \upsilon_{itj} \tag{1}$$

where the inclusion of γ_j distinguishes this partial treatment effect specification from classic DinD. Rearranging yields:

$$Y_{itj} = \mu + \tau D_{i1} + \delta t + \alpha D_{it} + X_{it}\beta + \gamma_j + \varepsilon_{itj} \tag{2}$$

where

$$\begin{aligned} \mu &= \mathbb{E}[\eta_i | D_{i1} = 0] + \delta_0 \\ \tau &= \mathbb{E}[\eta_i | D_{i1} = 1] - \mathbb{E}[\eta_i | D_{i1} = 0] \\ \delta &= \delta_1 - \delta_0 \\ \varepsilon_{itj} &= \eta_i - \mathbb{E}[\eta_i | D_{i1}] + \upsilon_{itj} \end{aligned} \tag{3}$$

For ease of notation, let $S_{it}(D_{i1}, t) = \{D_{i1}, t, X_{it}\}$ contain the relevant conditioning information of treatment group, period, and observed covariates. With this structure, expected grades conditional on X_{it} and γ_j before and after treatment for treated and non-treated students are given by:

$$\mathbb{E}\left[Y_{itj} | D_{i1} = 1, t = 1, X_{it}, \gamma_j\right] = \mu + \tau + \delta + \alpha + X_{it}\beta + \gamma_j + \mathbb{E}\left[\varepsilon_{itj} | S_{it}(1,1), \gamma_j\right] \tag{4}$$

$$\mathbb{E}\left[Y_{itj} | D_{i1} = 1, t = 0, X_{it}, \gamma_j\right] = \mu + \tau + X_{it}\beta + \gamma_j + \mathbb{E}\left[\varepsilon_{itj} | S_{it}(1,0), \gamma_j\right] \tag{5}$$

$$\mathbb{E}\left[Y_{itj} | D_{i1} = 0, t = 1, X_{it}, \gamma_j\right] = \mu + \delta + X_{it}\beta + \gamma_j + \mathbb{E}\left[\varepsilon_{itj} | S_{it}(0,1), \gamma_j\right] \tag{6}$$

$$\mathbb{E}\left[Y_{itj} | D_{i1} = 0, t = 0, X_{it}, \gamma_j\right] = \mu + X_{it}\beta + \gamma_j + \mathbb{E}\left[\varepsilon_{itj} | S_{it}(0,0), \gamma_j\right] \tag{7}$$

The double difference $(4-5)-(6-7)$ yields:

$$\begin{aligned}\alpha_{PTE} &= \alpha+\left(\mathbb{E}\left[\varepsilon_{itj}\,|\,S_{it}\,(1,1)\,,\gamma_j\right]-\mathbb{E}\left[\varepsilon_{itj}\,|\,S_{it}\,(1,0)\,,\gamma_j\right]\right) \\ &\quad -\left(\mathbb{E}\left[\varepsilon_{itj}\,|\,S_{it}\,(0,1)\,,\gamma_j\right]-\mathbb{E}\left[\varepsilon_{itj}\,|\,S_{it}\,(0,0)\,,\gamma_j\right]\right)\end{aligned} \tag{8}$$

The DinD identifying assumption is that idiosyncratic noise is mean independent of treatment status, that is:

$$\mathbb{E}\left[\upsilon_{itj}\,|\,D_{i1}\right]=\mathbb{E}\left[\upsilon_{itj}\right] \tag{9}$$

This implies the composite error ε_{itj} is mean independent of treatment status and period which yields $\alpha_{PTE}=\alpha$. Additionally, this assumption also implies unbiased estimates of parameters in (2) can be obtained using OLS. Together, this shows DinD estimates of the partial effect of Greek affiliation on academic outcomes can be obtained using panel or repeated cross-sectional data by regressing observed grades on a constant, a dummy variable for ever affiliating with a Greek organization, a dummy variable for post-treatment period, a dummy variable for being affiliated in this period, observed characteristics, and course specific dummy variables.

5.3 Total Treatment Effect

In this subsection, we present a DinD approach to identify the total treatment effect which allows students to re-optimize their course choice behavior after joining a Greek organization. In the previous subsection, we describe an estimator which explicitly controls for these behavioral changes; however, there may be applications where this is not desirable. Students themselves may actually prefer to know the change in grade point average (GPA) they should expect to experience from Greek affiliation allowing for the changes in course choice behavior which typically result from affiliation. Additionally, we show the difference between the total effect and average effect represents the effect of Greek affiliation on the grading leniency of chosen courses.

To show identification of the total effect, use Equation (2) to derive expected grades conditional on X_{it} (but not γ_j) before and after treatment for treated and non-treated students:

$$\mathbb{E}\left[Y_{itj}\,|\,D_{i1}=1, t=1, X_{it}\right] = \mu+\tau+\delta+\alpha+X_{it}\beta+\mathbb{E}\left[\gamma_j\,|\,S_{it}(1,1)\right]+\mathbb{E}\left[\varepsilon_{itj}\,|\,S_{it}(1,1)\right] \quad (10)$$

$$\mathbb{E}\left[Y_{itj}\,|\,D_{i1}=1, t=0, X_{it}\right] = \mu+\tau+X_{it}\beta+\mathbb{E}\left[\gamma_j\,|\,S_{it}(1,0)\right]+\mathbb{E}\left[\varepsilon_{itj}\,|\,S_{it}(1,0)\right] \quad (11)$$

$$\mathbb{E}\left[Y_{itj}\,|\,D_{i1}=0, t=1, X_{it}\right] = \mu+\delta+X_{it}\beta+\mathbb{E}\left[\gamma_j\,|\,S_{it}(0,1)\right]+\mathbb{E}\left[\varepsilon_{itj}\,|\,S_{it}(0,1)\right] \quad (12)$$

$$\mathbb{E}\left[Y_{itj}\,|\,D_{i1}=0, t=0, X_{it}\right] = \mu+X_{it}\beta+\mathbb{E}\left[\gamma_j\,|\,S_{it}(0,0)\right]+\mathbb{E}\left[\varepsilon_{itj}\,|\,S_{it}(0,0)\right] \quad (13)$$

In this case, the double difference $(10-11)-(12-13)$ yields:

$$\begin{aligned}\alpha_{TTE} = {} & \alpha+\left(\mathbb{E}\left[\gamma_j\,|\,S_{it}(1,1)\right]-\mathbb{E}\left[\gamma_j\,|\,S_{it}(1,0)\right]\right) \\ & -\left(\mathbb{E}\left[\gamma_j\,|\,S_{it}(0,1)\right]-\mathbb{E}\left[\gamma_j\,|\,S_{it}(0,0)\right]\right) \\ & +\left(\mathbb{E}\left[\varepsilon_{itj}\,|\,S_{it}(1,1)\right]-\mathbb{E}\left[\varepsilon_{itj}\,|\,S_{it}(1,0)\right]\right) \\ & -\left(\mathbb{E}\left[\varepsilon_{itj}\,|\,S_{it}(0,1)\right]-\mathbb{E}\left[\varepsilon_{itj}\,|\,S_{it}(0,0)\right]\right)\end{aligned} \quad (14)$$

Notice this expression contains the expected change in grading leniency γ_j conditional on X_{it} for Greek students before and after affiliation relative to the expected change in grading leniency over the same period for independents. Let Δ represent this difference-in-differences:

$$\Delta = \left(\mathbb{E}\left[\gamma_j\,|\,S_{it}(1,1)\right]-\mathbb{E}\left[\gamma_j\,|\,S_{it}(1,0)\right]\right)-\left(\mathbb{E}\left[\gamma_j\,|\,S_{it}(0,1)\right]-\mathbb{E}\left[\gamma_j\,|\,S_{it}(0,0)\right]\right) \quad (15)$$

Under the same identifying assumption given in Equation (9):

$$\alpha_{TTE} = \alpha+\Delta \quad (16)$$

This implies omitting course specific dummy variables from OLS estimation of Equation (2) biases estimates of α by exactly Δ. Under an assumption analogous to (9), Δ identifies the effect of Greek affiliation on grading leniency. This implies the difference between the total treatment effect and partial treatment effect can be used to identify the effect of Greek affiliation on the grading leniency of chosen courses.

In Section 7, we report estimates of the partial treatment effect, total treatment effect, and leniency effect for every specification. This provides a multidimensional analysis of the effect of Greek affiliation on academic outcomes. The partial treatment effect measures pure change in academic performance, the total treatment effect measures the change in expected outcomes, and the leniency effect measures how Greek affiliation affects grading leniency of chosen courses.

5.4 Censoring at 4.0

In this subsection, we discuss DinD estimation and interpretation when the dependent variable is censored. To our knowledge, this is the first paper to provide such a discussion. These tools may be applied to estimate treatment effects in a variety of settings with censored dependent variables such as labor market outcomes (Chay and Powell, 2001) or health outcomes (Austin, Escobar, and Kopec, 2000). We show that standard techniques for estimating censored regression models may be applied to a DinD framework with important caveats regarding interpretation and the impact of serial correlation in errors.

If grades are censored at 4.0, Equation (2) should be adjusted to:

$$\begin{aligned} Y^{\star}_{itj} &= \mu + \tau D_{i1} + \delta t + \alpha D_{it} + X_{it}\beta + \gamma_j + \varepsilon_{itj} \\ Y_{itj} &= \min\left\{Y^{\star}_{itj}, 4\right\} \end{aligned} \tag{17}$$

where $Y^{\star}_{itj}$ represents latent (unobserved) grades and Y_{itj} represents observed grades. μ, τ, δ, and ε_{itj} are as defined in Equation (3). To simplify notation, let $\tilde{Y}^{\star}_{itj} = \mu + \tau D_{i1} + \delta t + \alpha D_{it} + X_{it}\beta + \gamma_j$ denote the deterministic component of latent grades.

Censoring confounds DinD estimators because positive (negative) treatment effects increase (decrease) the probability of observations being censored. To illustrate this, let

$$P(D_{i1},t) = \Pr\left(Y^{\star}_{itj} \leq 4 \,|\, S_{it}(D_{i1},t), \gamma_j\right) \tag{18}$$

be the conditional probability that an earned grade is uncensored. Conditional expectations of observed grades are then given by:

$$\begin{aligned} \mathbb{E}\left[Y_{itj} \,|\, S_{it}(D_{i1},t), \gamma_j\right] &= 4\left[1 - P(D_{i1},t)\right] + \mathbb{E}\left[Y^{\star}_{itj} \,|\, S_{it}(D_{i1},t), \gamma_j\right] P(D_{i1},t) & (19) \\ &= 4\left[1 - P(D_{i1},t)\right] \\ &\quad + \left\{\tilde{Y}^{\star}_{itj} + \mathbb{E}\left[\varepsilon_{itj} \,|\, S_{it}(D_{i1},t), \varepsilon_{itj} \leq \left(4 - \tilde{Y}^{\star}_{itj}\right)\right]\right\} P(D_{i1},t) & (20) \end{aligned}$$

Except in the trivial case when there are no treatment effects, selection effects, or period effects, $P(D_{i1},t)$ varies with D_{it} and t. As such, the double difference will be a complicated function of $P(D_{i1},t)$; additionally, confounding factors δ_t, $X_{it}\beta$, γ_j, and $\mathbb{E}\left[\varepsilon_{itj} \,|\, S_{it}(D_{i1},t), \varepsilon_{itj} \leq \left(4 - \tilde{Y}^{\star}_{itj}\right)\right]$ will not cancel. As such, DinD estimators typically will not identify treatment effects in the presence of censoring.

To identify treatment effects in the presence of censoring, we use a Maximum Likelihood approach which accounts for the effects of treatment, selection, and period on the probability that an earned grade is uncensored. This approach is an adaption of classical techniques for incorporating censoring (Tobin, 1958; Ameniya, 1984).

Let $F_{\varepsilon_{itj}}(\cdot)$ and $f_{\varepsilon_{itj}}(\cdot)$ represent the CDF and PDF for ε_{itj} and let $\Theta = \{\mu, \tau, \delta, \alpha, \beta, \gamma\}$ contain all parameters.[15] The conditional likelihood of observing Y_{itj} is then given by:

$$\mathscr{L}_{itj}\left(Y_{itj} \,|\, S_{it}(D_{i1},t), \Theta\right) = \left[1 - F_{\varepsilon_{itj}}\left(4 - \tilde{Y}^{\star}_{itj}\right)\right]^{(Y_{itj}=4)} f_{\varepsilon_{itj}}\left(Y_{itj} - \tilde{Y}^{\star}_{itj}\right)^{(Y_{itj}<4)} \tag{21}$$

The component $\left[1 - F_{\varepsilon_{itj}}\left(4 - \tilde{Y}^{\star}_{itj}\right)\right]^{(Y_{itj}=4)}$ measures the probability that grades which are cen-

[15] Estimation requires the econometrician to observe $F_{\varepsilon_{itj}}(\cdot)$ and $f_{\varepsilon_{itj}}(\cdot)$. We follow standard practices and assume $\varepsilon_{itj} \sim N\left(0, \sigma^2\right)$ (Hayashi, 2000 Ch. 8).

sored in the data are also censored in the model. The component $f_{\varepsilon_{itj}}\left(Y_{itj}-\tilde{Y}^{\star}_{itj}\right)^{(Y_{itj}<4)}$ measures the likelihood that uncensored grades in the data have the same values as in the model. As such, $\mathscr{L}_{itj}\left(Y_{itj}\,|\,S_{it}\left(D_{i1},t\right),\Theta\right)$ includes the likelihood of observing all observed censored and uncensored grades in the data.

If ε_{itj} were iid, the full log-likelihood would be given by:

$$\ln\mathscr{L}=\sum_{i=1}^{N}\sum_{t=0}^{1}\left(Y_{itj}=4\right)\ln\left[1-F_{\varepsilon_{itj}}\left(4-\tilde{Y}^{\star}_{itj}\right)\right]+\left(Y_{itj}<4\right)\ln\left[f_{\varepsilon_{itj}}\left(Y_{itj}-\tilde{Y}^{\star}_{itj}\right)\right] \tag{22}$$

In this setting, ε_{itj} is likely to be serially correlated because it includes factors which are fixed within individuals. However, Hayashi (2000, Ch 7) shows misspecified maximum likelihood estimators—or ‘quasi-ML estimators’—are consistent under various forms of misspecification including erroneously treating serially correlated observations as iid. Robinson (1982) explicitly considers this type of misspecification in censored regression models and shows the quasi-ML estimator which maximizes (22) yields consistent (but not efficient) estimates. This shows classical censored regression estimation techniques can be used to consistently (but not efficiently) estimate the parameters of equation (17).

Estimates of the parameters of equation (17) can be used to construct treatment effects on either latent outcomes $Y^{\star}_{itj}$ or observed outcomes Y_{itj}. Treatment effects on latent outcomes are directly obtained from estimates of α. In this simplified setting, these treatment effects are the same for all individuals. Following Puhani (2012), treatment effects on observed outcomes can be obtained by using (17) to take the difference between expected observed grades Y_{itj} in treated and untreated states for every observation. In general, censoring causes these treatment effects to differ across observations. In our context, we feel the effects on latent outcomes are more relevant because they measure the effects on true academic performance. As such, we report treatment effects on latent outcomes in Section 7.

6 Empirical Specification

In this section, we present the specific empirical specifications used in our analysis and discuss the plausibility of the identifying assumptions for our difference-in-differences (DinD) approach. We begin by presenting our baseline empirical specification which estimates gender specific and time varying treatment effects. Next we describe empirical specifications which also allow treatment effects to vary based on the initial preparation of the student and the social status of the Greek organization. Following this we discuss the plausibility of the common trends assumption in the case of both the effect of Greek affiliation on academic performance and the effect of Greek affiliation on grading leniency. Finally, we discuss the plausibility of the compositional consistency assumption.

6.1 Gender Specific Time Varying Treatment Effects

In this subsection, we present our empirical specification for estimating partial and total treatment effects which vary by gender and time period. Index students by $i = 1, \dots, N$, student gender by $g \in \{M, F\}$, courses by $j = 1, \dots, J$, and academic semesters by $t = 1, \dots, 7$.[16] Let $D_i = 1$ if student i ever affiliates with a Greek organization. Our specification for estimating partial treatment effects is given by:

$$\begin{aligned} Y^{\star}_{itj} &= \mu_g + \tau_g D_i + \sum_{t'=2}^{7} \delta_{t'g} \left(t = t'\right) + \alpha_{2g} D_i \left(t = 2\right) + \sum_{t' \in \{3,6\}} \alpha_{fg} D_i \left(t = t'\right) \\ &\quad + \sum_{t' \in \{4,5,7\}} \alpha_{sg} D_i \left(t = t'\right) + X_i \Gamma_t + Z_i \Phi + \gamma_j + \varepsilon_{itj} \qquad (23) \\ Y_{itj} &= \min\left\{Y^{\star}_{itj}, 4\right\} \end{aligned}$$

[16] Academic semesters are given by Freshman Fall - 1, Freshman Spring - 2, Sophomore Fall - 3, Sophomore Spring - 4, Junior Spring - 5, Senior Fall - 6, Senior Spring - 7. We exclude Junior Fall because a large portion of the sample studies abroad causing selection concerns. Summer terms are excluded for the same reason. Courses are defined by course number and academic term. For statistical power, courses with fewer than 10 students are pooled into groups defined by academic subjects and course level.

μ_g is a gender specific intercept, τ_g is a gender specific parameter measuring unobserved differences between the Greek and independent populations, δ_{tg} are gender specific time trends, α_{tg} are gender specific treatment parameters, X_i are baseline covariates with time varying effects, Z_i are baseline covariates with time invariant effects, and γ_j is a course fixed effect capturing grading leniency.[17] As discussed in Section 5, our specification for estimating total treatment effects modifies (23) by omitting course fixed effects γ_j. Furthermore, the effects of Greek affiliation on grading leniency are estimated by taking differences between total treatment effects and partial treatment effects.[18]

This specification includes three treatment effects for each gender: the effect in Freshman Spring (α_{2g}), the effect in Fall semesters after Freshman year (α_{fg}), and the effect in Spring semesters after Freshman year (α_{sg}). This allows for separate treatment effects in the semester when students undergo new member education and in subsequent years depending on whether the organization is recruiting and educating new members. To flexibly account for gender heterogeneity we allow all DinD parameters μ_g, τ_g, δ_{tg}, and α_{tg} to vary by gender.

6.2 Heterogeneous Effects by Initial Preparation

To further understand how Greek affiliation affects various members of the Duke population we also estimate treatment effects which vary by the student's composite SAT scores. Intuitively, this involves defining four treatment groups—low SAT fraternity members, high SAT fraternity members, low SAT sorority members, and high SAT sorority members—and four control groups of analogous independent students. Changes in performance for each treatment group are com-

[17] We report all estimates for three possible specifications of X_i and Z_i. As a baseline, the first specification includes no controls. The second specification includes detailed demographic data in X_i but no controls in Z_i. As we discuss in further detail later in this section, we allow demographic data to have time varying effects to strengthen the plausibility of the common trends assumption. Our third specification adds composite SAT scores and five internal measures of application quality (academic achievement, high school curriculum difficulty, essay quality, personal qualities, and letters of recommendation) to Z_i. Adding these baseline preparation controls improves the precision of estimates.

[18] Standard errors for the effects of Greek affiliation on grading leniency are estimated using a block bootstrap algorithm as follows: repeatedly sample N students with replacement; estimate α_{ATTT} and α_{APTT} for each sample; compute the difference $\hat{\Delta} = \hat{\alpha}_{ATTT} - \hat{\alpha}_{APTT}$ for each sample; estimate the standard deviation in the distribution of $\hat{\Delta}$ across samples.

pared to changes in performance for the corresponding control group to flexibly estimate separate treatment effect parameters for more and less prepared male students as well as more and less prepared female students. We divide students into preparation levels $p \in \{Low, High\}$ based on whether their composite SAT score falls below or above the sample median.[19] Our specification for estimating partial treatment effects is given by:

$$\begin{aligned} Y^{\star}_{itj} &= \mu_{gp} + \tau_{gp} D_i + \sum_{t'=2}^{7} \delta_{t'gp}\left(t = t'\right) + \alpha_{2gp} D_i \left(t = 2\right) + \sum_{t' \in \{3,6\}} \alpha_{fgp} D_i \left(t = t'\right) \\ &\quad + \sum_{t' \in \{4,5,7\}} \alpha_{sgp} D_i \left(t = t'\right) + X_i \Gamma_t + Z_i \Phi + \gamma_j + \varepsilon_{itj} \qquad (24) \\ Y_{itj} &= \min\left\{Y^{\star}_{itj}, 4\right\} \end{aligned}$$

To estimate separate treatment effects for every possible gender-preparation pair, all DinD parameters μ_{gp}, τ_{gp}, δ_{tgp}, and α_{tgp} vary by gender-preparation pairs. As before, total treatment effects are estimated by omitting course fixed effects γ_j and effects on grading leniency are estimated by the difference between total treatment effects and partial treatment effects.

6.3 Heterogeneous Effects by Social status

As discussed in Section 2, there are well known social hierarchies within fraternities and sororities. In this subsection, we present our empirical specification for estimating partial and total treatment effects which also depend on the social status of the organization. In this specification, there are two potential treatment levels for both male and female students: high status organizations and low status organizations. Both high status and low status treatment groups are compared to independents of the same gender to flexibly estimate treatment effects for high status and low status sororities and fraternities. As discussed in Section 2, we divide Greek organizations into social tiers $r \in \{Low, High\}$ based on a informal perusal of postings on the forum Greekrank.com.

[19] The sample median is 1420 out of 1600. Students who are missing SAT scores are categorized by whether the average of their internal admissions scores is above or below the sample median.

Our specification for estimating partial treatment effects is given by:

$$
\begin{aligned}
Y^{\star}_{itj} &= \mu_g + \tau_{gr} D_i + \sum_{t'=2}^{7} \delta_{t'g} \left(t = t'\right) + \alpha_{2gr} D_i \left(t = 2\right) + \sum_{t' \in \{3,6\}} \alpha_{fgr} D_i \left(t = t'\right) \\
&\quad + \sum_{t' \in \{4,5,7\}} \alpha_{sgr} D_i \left(t = t'\right) + X_i \Gamma_t + Z_i \Phi + \gamma_j + \varepsilon_{itj} \qquad (25) \\
Y_{itj} &= \min \left\{ Y^{\star}_{itj}, 4 \right\}
\end{aligned}
$$

To accommodate multiple treatment levels, the treatment and selection parameters τ_{gr} and α_{tgr} vary by social status and gender while the control group parameters μ_g and δ_{tg} vary by gender only. As before, total treatment effects are estimated by omitting course fixed effects γ_j and effects on grading leniency are estimated by the difference between total treatment effects and partial treatment effects.

6.4 Threats to Identification: Common trends assumption for the effects of Greek affiliation on academic performance

A crucial identifying assumption for DinD estimators is the common trends assumption which states that students who join a Greek organization would have experienced the same change in performance over time as independents in a hypothetical scenario when they never join Greek organizations (Blundell and Costa Dias, 2009). This is necessary to make independent students an appropriate comparison group for Greek students.

There are several reasons to be concerned about the common trends assumption. First, we may be concerned that differences between the Greek and independent populations may have differential impacts over time. Section 4 shows Greek students generally come from wealthier and more educated families. It is possible that family background has a differential effect on academic performance over time which would violate the common trends assumption. To address this, we follow Duflo (2001) and estimate a specification which includes detailed demographic data in X_i allowing the effect of these variables to vary over time.[20] This allows observed demographic

[20] X_i includes indicator variables for education of the most educated parent (less than Bachelors degree, Bachelor's

variables such as family background and race to have differential effects over time which strengthens the plausibility of the common trends assumption. Furthermore, in Section 7 we show these controls have negligible effects on treatment effect estimates which supports the common trends assumption.

Second, we may be concerned that decisions to join Greek organizations are somehow related to the unobserved component of performance in the pre-treatment Freshman Spring semester. This may happen if the decision to affiliate with a Greek organization depends on academic performance in the Freshman Fall. It may also happen if students who plan to join Greek organizations work especially hard in the Freshman Fall semester to compensate for anticipated negative effects of affiliating. This is akin to the well-known 'Ashenfelter's dip' of program evaluation literature which violates the common trend assumption through mean reversion (Ashenfelter, 1978). Unfortunately, we are unable to control for this complication. If the probability of affiliating is positively (negatively) correlated with the unobserved component of grade production in the Freshman Fall semester then we will over-state (under-state) the negative effects of affiliating.

6.5 Threats to Identification: Common trends assumption for the effects of Greek affiliation on grading leniency

In Section 5, we showed that under the identifying assumption (9), the difference between the total treatment effect and the partial treatment effect measures the expected change in chosen grading leniency conditional on X_{it} for Greek students before and after affiliation relative to the expected change in chosen grading leniency over the same period for independents.

$$\begin{aligned} \Delta \;=\; & \left(\mathbb{E}\left[\gamma_j \,|\, D_{i1}=1, t=1, X_{it}\right] - \mathbb{E}\left[\gamma_j \,|\, D_{i1}=1, t=0, X_{it}\right]\right) \\ & - \left(\mathbb{E}\left[\gamma_j \,|\, D_{i1}=0, t=1, X_{it}\right] - \mathbb{E}\left[\gamma_j \,|\, D_{i1}=0, t=0, X_{it}\right]\right) \end{aligned} \tag{26}$$

or Master's degree, Professional degree); family income (less than \$50,000 per year, \$50,000 - \$99,999 per year, \$100,000 - \$199,999 per year, or \$200,000 or more per year); student race (White, Black, Hispanic, Asian, Other Race); whether the student is a legacy; and whether the student attended a private high school.

As such, this difference is itself a conditional difference-in-differences estimator of the effects of Greek affiliation on grading leniency γ_j.

For this estimator, the common trends assumption states that students who join a Greek organization would have experienced the same change in grading leniency over time as independents in a hypothetical scenario when they never join Greek organizations. The reasons to be concerned about the common trends assumption for this estimator are similar to the concerns discussed in the preceding subsection. First, demographic characteristics may affect the time path of grading leniency; for example, students from more educated families may be more willing to take harshly graded courses at the beginning of their college career. As before, allowing rich demographic variables to have differential effects over time strengthens the plausibility of this assumption. Second, we may also be concerned that students who plan to affiliate take lenient courses in the Freshman Fall semester to provide a cushion against the negative effects of affiliating in the following semester. Once again, we cannot control for this Ashenfelter's dip; however, university and major requirements leave students limited scope for strategic course choices in the Freshman Fall semester ("Course Selection," 2015). If Greek students choose unusually lenient (harsh) courses in the Freshman Fall then we will understate (overstate) the tendency of Greek organizations to draw their members into less challenging courses.

6.6 Threats to Identification: Compositional consistency assumption

The other crucial identifying assumptions for DinD estimators is the compositional consistency assumption which states that the composition of the Greek and independent samples must be the same in all periods (Blundell and Costa Dias, 2009). If either sample changes selectively over time then DinD estimators will confound treatment effects with these compositional changes.

We make several considerations to ensure the compositional consistency assumption is satisfied. First, although dropping out of Duke University is quite rare, there are several students that are not present in all eight Fall and Spring academic semesters. The most common reason for these missing semesters is students studying abroad in the Fall semester of Junior year. To address this we exclude the Junior Fall semester from our analysis. Students may be missing in other semesters if they graduate early, study abroad in a non-standard semester, take a leave of absence, or leave Duke altogether. To ensure sample consistency, we drop 124 students who have no course observations in at least one semester other than Junior Fall. In an online appendix, we show this does not change the representativeness of our sample. Second, sample composition may be altered by students taking more or fewer courses in any given semester. Intuitively, if all students take similar course loads in Freshman Fall but ambitious students take heavier course loads in subsequent semesters student-course level data will over-represent ambitious students in these later semesters. One solution to this problem is to use student-term specific averages rather than student-course data; however, these averages would not properly account for censoring in grades. As an alternative solution, we weight observations of student i in academic semester t by:

$$w_{it} = \frac{\omega_i}{n_{it}} \left(\frac{\sum_{i'=1}^{N} \sum_{t'=1}^{7} n_{i't'}}{\sum_{i'=1}^{N} \sum_{t'=1}^{7} \left(\frac{\omega_{i'}}{n_{i't'}} \right)} \right) \tag{27}$$

where n_{it} is the number of courses taken by student i in term t and ω_i are survey weights on students. These observation weights give more (less) weight to student-course observations from semesters where the student took a light (heavy) course load; as such, they ensure that every student receives the same weight in every semester regardless of their course load.

7 Results

In this section, we discuss estimation results which measure the effects of Greek affiliation on academic performance. We begin by discussing estimates of time varying effects for both fraternities and sororities. Next we compare these results to naive estimates which ignore the confounding effects of censoring at 4.0. Following this, we discuss estimates which also allow treatment effects to vary by a student's initial preparation. Finally, we present estimates which allow treatment effects to differ based on the social status of the Greek organization.

Each analysis reports total effects of affiliation on academic performance, partial effects of affiliation on performance, and effects of affiliation on grading leniency. As discussed in Section 5, partial effects control for changes in course choice behavior possibly resulting from affiliation while total effects allow students to re-optimize how they choose courses following affiliation. Each analysis also includes three specifications for controlling variables: first, a baseline specification with no controls; second, a specification which allows demographic variables to have time varying effects to strengthen the plausibility of the common trends assumption; third, a specification which adds measures of baseline preparation to increase precision. Finally, each analysis also includes estimates of selection effects τ_g to assess whether students who affiliate with Greek organizations are unobservably stronger or weaker than their independent peers. Standard errors are estimated using block bootstrap algorithms to account for the inherent serial correlation in DinD errors (Bertrand, Duflo, and Mullainathan, 2004).

7.1 Gender Specific Time Varying Treatment Effects

Table 2 contains estimates which allow effects to vary over time and by gender. To understand how recruitment and new member education affect academic outcomes we report separate treatment effects for the Freshman Spring semester (when members affiliate), Fall semesters after Freshman year (when limited recruitment occurs), and Spring semesters after Freshman year (when students may recruit and educate new members).

The fraternity panel of results suggest the process of joining a fraternity is a significant distraction from coursework. In the Freshman Spring semester, the total effect of joining a fraternity is -.168 grade points. For reference, this is 24.7% of a standard deviation in the grade distribution and is approximately half as large as the difference between a B and a B+ (.3 grade points). This total effect also ignores the fact that fraternity members strategically choose leniently graded courses in this semester to mitigate the negative effects of affiliation. If one controls for these behavioral changes, the partial effect of joining a fraternity is -.216 grade points in the Freshman Spring semester. As discussed in Section 5, the difference between the total effect and partial effect measures the effect of affiliation on grading leniency. The results suggest male students respond to the burdens of fraternity recruitment and new member education by choosing courses which are .048 grade points more lenient in their grading. This mitigates the negative effects of affiliation by 22.2%.

The fraternity panel of results also suggests fraternity affiliation has modest negative effects in semesters after Freshman year. In Fall semesters after Freshman year, the total effect is -.099 grade points while the partial effect is -.132 grade points implying behavioral responses mitigate the negative effects of affiliation by 24.7%. In Spring semesters after Freshman year, the total effect is insignificant while the partial effect is -.089 grade points.

Finally, the 'Selection' row of the fraternity panel of results reports estimates of τ_g which measure whether fraternity members are unobservably stronger or weaker students than their independent peers. The results show there is modest positive selection into fraternities which becomes small and insignificant when rich controls for baseline preparation are included. This implies rich demographic controls and measures of baseline preparation account for most of the confounding factors which affect selection into fraternities.

The sorority panel of results shows a very different pattern: for sorority members, the worst distraction occurs in Spring semesters after Freshman year when students are recruiting and educating new members. In these semesters, the total effect of joining a sorority is -.109 grade points or 16.0% of a standard deviation. Once again, this understates the true effect because sorority members choose more lenient classes to reduce the effects of affiliation. Controlling for changes in course choice behavior yields a partial effect of -.146 grade points. Sorority affiliation also has modest negative effects in the Freshman Spring semester and in Fall semesters after Freshman year; however, the effects are almost completely eliminated by choosing courses which grade more leniently. In the Freshman Spring semester, these responses reduce negative effects by 29.3% while in Fall semesters after Freshman Spring these responses mitigate effects by 64.5%.

Interestingly, the selection effect estimates for sorority members suggest there is strong positive selection into sororities which is only partially eliminated with rich controls for demographic information and baseline preparation. With full controls, sorority members are unobservably stronger by .122 grade points after accounting for differences in grading leniency. This demonstrates the importance of controlling for selection on unobserved factors: the selection effect is approximately the opposite of the treatment effects—ignoring selection on unobserved factors would yield estimated effects close to zero.

Finally, we can compare treatment effect estimates in 'No Controls' columns to those in 'Demographics' columns to assess the sensitivity of estimates to time varying effects of demographics. As discussed in Section 4, Greek students are substantially whiter and wealthier than independents raising concerns about the common trends assumption of DinD. The 'Demographics' column of results includes categorical controls for parent's education, family income, race, high school type, and legacy status which are allowed to have different effects across semesters. We see that including these time varying demographic controls has negligible effects on treatment effect estimates which suggests observed baseline differences between Greeks and independents are not confounding our estimates.

7.2 Ignoring Censoring

Table 3 quantifies the confounding effects of censoring at 4.0 by reporting the same estimates as Table 2 using OLS. Results show ignoring censoring leads to understating the negative effects of Greek affiliation. For fraternity members, the total effect of affiliating in Freshman Spring is understated by 22.0% while the partial effect of affiliating in Freshman Spring is understated by 19.4%. For sorority members, the total effect of affiliating in Spring semesters after Freshman year is understated by 22.9% while the partial effect of affiliating in Spring semesters after Freshman year is understated by 19.2%.

Intuitively, this substantial attenuation bias occurs because more able Greek students earn 4.0 grade points in both their observed state and in the counter-factual scenario in which they remain independent. For these students, treatment effects appear to be zero. Averaging these students with the rest of the population yields attenuated estimates of the negative effects of affiliation.

These results imply it is important to account for censoring at 4.0 when estimating effects on academic outcomes in higher education. Substantial grade inflation at Duke and at universities across the US right censor the academic outcomes of bright (and sometimes even mediocre) students (Mansfield, 2001). As we show, this censoring can have strong attenuating effects on treatment effect estimates.

7.3 Heterogeneous Effects by Initial Preparation

Tables 4 and 5 extend the analysis of Table 2 by also allowing treatment effects to vary based on the students initial preparation. We divide students into preparation levels $p \in \{Low, High\}$ based on whether their composite SAT score falls below or above the sample median and define four treatment groups and four control groups by gender and preparation level.[21] Intuitively, we compare the change in performance for each treatment group before and after affiliation to the change in performance of the corresponding control group over the same period.

The low SAT panel of results in Table 4 show fraternities have especially negative effects on less prepared students. In the Freshman Spring semester, the total effect of joining a fraternity on low SAT students is -.320 grade points. This is a very large effect: -.320 grade points represents 47.1% of a standard deviation in the grade distribution and is slightly more than the difference between a B and a B+ (.3 grade points). Furthermore, the large negative effects of fraternities on low SAT students do not stop after Freshman year: in Fall semesters after Freshman year the total effect is -.206 grade points and in Spring semesters after Freshman year the total effect is -.126 grade points.

Interestingly, low SAT fraternity members do not respond to these negative effects by choosing courses which grade more leniently. Two theories could explain why these students are not responding to these strong negative effects: First, it may be that male students with low SAT scores already take lenient courses leaving limited scope for switching to more lenient classes. Second, it

[21] The sample median is 1420 out of 1600. Students who are missing SAT scores are categorized by whether the average of their internal admissions scores is above or below the sample median.

may be that male students with low SAT scores are less concerned with grades and thus unwilling to change their course choice behavior to improve their GPA. Either way, the limited responses imply the negative effects in terms of observed outcomes and in terms of true performance are both worryingly large.

Finally, estimates of selection effects for low SAT fraternity members suggest there is strong positive selection into fraternities among low SAT male students. These selection effects persist even after including rich controls for demographic information and baseline preparation. With full controls, low SAT fraternity members are unobservably stronger by .216 grade points after accounting for differences in grading leniency. Once again, this demonstrates the importance of controlling for selection on unobserved factors: a study which ignored this selection would drastically understate the negative effects of fraternity affiliation on less prepared students.

The high SAT panel of results in Table 4 show fraternities have very different effects on more prepared students. These students are much less affected than their less prepared peers: In semesters after Freshman year, affiliation has small and insignificant effects on these students. In the Freshman Spring semester, there is a partial effect of -.120 grade points but this is mitigated by 60.7% as a result of choosing more lenient courses.

This shows fraternities affect the academic outcomes of their members in very different ways. More prepared students are burdened by new member education but are able to prevent any harm to their GPA by choosing lenient courses in this semester. After Freshman year, more prepared students are generally unaffected. Conversely, less prepared students experience strong negative effects throughout their college career and are unwilling (or unable) to react to the effects by choosing more lenient courses.

The results for female students in Table 5 show sororities also have larger effects on less prepared students. In Spring semesters after Freshman year (the worst semesters for sororities), the total effect of joining a sorority on low SAT students is -.143 grade points or 21.0% of a standard deviation. Unlike low SAT male students, low SAT female students do respond to these negative effects by choosing more lenient courses. After controlling for these responses, the partial effect of joining a sorority on low SAT students in Spring semesters after Freshman year is -.178 grade points. The behavioral responses of low SAT sorority members are especially strong in Fall semesters after Freshman year. In these semesters, low SAT sorority members choose courses which are .071 grade points more lenient reducing the negative effects of affiliation by 49.3%.

As with low SAT men, estimates of selection effects for low SAT women show there is strong positive selection into sororities among low SAT female students. Once again, these differences cannot be explained by demographic characteristics or initial preparation and these selection effects are similar in magnitude to the treatment effects. This shows it is especially important to control for selection on unobserved factors when estimating the effects of affiliation on less prepared students.

The high SAT panel of results in Table 5 show sororities also have much smaller effects on more prepared students. All estimated effects are small and statistically insignificant. As with fraternities, sorority life is a substantial distraction for less prepared students but has limited effects on students who are more prepared.

These heterogeneous effects by initial preparation have important implications for policy: First, many universities—including Duke—publish the average GPA of different Greek organizations to inform students and parents and to reward high performing organizations ("Academic Data", 2015). However, these averages ignore heterogeneous effects across students. Universities could define 'at risk students' by SAT scores or other criteria and report separate averages for these students to more carefully monitor the experiences of less prepared students who are generally more affected by Greek affiliation. More drastically, universities could make Greek eligibility contingent on academic performance to prevent vulnerable students from ever affiliating.

7.4 Heterogeneous Effects by Social Status

To further examine heterogeneous effects of affiliation, Tables 6 and 7 reports separate treatment effects by the social status of the Greek organization. As discussed in Section 2, there are well known social hierarchies within fraternities and sororities. We divide Greek organizations into social tiers $r \in \{Low, High\}$ based on an informal perusal of postings on the forum Greekrank.com.

Intuitively, we estimate separate treatment effects by comparing the change in performance for members of high status and low status Greek organizations before and after affiliation to the change in performance over the same period for independents. We note that social status is mean independent of whether a student has high or low SAT scores; as such, the heterogeneity discussed in this subsection arises from different variation than the heterogeneity discussed in the preceding section.[22]

The low status panel of results in Table 6 show low status fraternities have strong negative effects in both the Freshman Spring semester and in Fall semesters after Freshman year. In the Freshman Spring semester, the partial effect of joining a low status fraternity is -.172 grade points or 25.3% of a standard deviation. In Fall semesters after Freshman year, the partial effect of joining a low status fraternity is -.160 grade points or 23.5% of a standard deviation. These strong negative effects are somewhat mitigated by behavioral responses which reduce these partial effects by 16.3% and 17.5% respectively. Interestingly, estimates of selection effects suggest there is very little confounding selection into low status fraternities. Both with and without controls, low status fraternity members are very similar to male independents.

The high status panel of results in Table 6 show high status fraternities are more distracting during new member education but less burdensome afterwards. In the Freshman Spring semester, the partial effect of joining a high status fraternity is -.254 grade points or 37.4% of a standard deviation. To handle the distractions of new member education, students joining high status fraternities choose courses which are .064 grade points more lenient which reduces the negative effects by 25.2%. In subsequent semesters, high status fraternities have small and insignificant effects on the academic performance of their members. Furthermore, unlike low status fraternities, estimates of selection effects suggest there is strong positive selection into high status fraternities which cannot be explained by demographic characteristics or baseline preparation.

These heterogeneous results paint an interesting picture of fraternity life at Duke. Fraternities higher up in the social hierarchy recruit better students, have new member education that causes a

[22] 49% of members of high status Greek organizations have SAT scores above the sample median while 48% of members of low status Greek organizations have SAT scores above the sample median.

substantial distraction from coursework, and have limited effects after Freshman year. Conversely, fraternities lower down in the social hierarchy recruit average students, have new member education which is less taxing, but cause more of a distraction in later years. It seems unsurprisingly that higher status fraternities have more demanding new member education: senior members of these fraternities likely feel a stronger obligation to uphold high standards and are less worried that new members will defect. We find it somewhat surprising that low status organizations are more of a distraction after Freshman year. One possible explanation for this is that full membership in a high status fraternity increases self esteem which mitigates the distracting effects of fraternity affiliation. Another possible explanation is that high status fraternities care more about their academic reputation than low status fraternities.

The results for female students in table 7 show high status sororities also shift more of their burden to new member education. Low status sororities have small and insignificant effects during the Freshman Spring semester but large effects in subsequent semesters. In Spring semesters after Freshman year (the worst semesters for sororities), the partial effect of joining a low status sorority is -.176 grade points or 25.9% of a standard deviation. Effects are also large in Fall semesters after Freshman year: in these semesters, the partial effect of joining a low status sorority is -.128 grade points or 18.8% of a standard deviation. These negative effects are mitigated by large behavioral responses which reduce effects by 32.4% and 46.7% respectively. Unlike low status fraternities, estimates of selection effects suggest there is strong positive selection into low status sororities which cannot be explained by demographic characteristics or baseline preparation.

The high status panel of results in Table 7 show high status sororities have more demanding new member education but are less distracting in later semesters. Spring semesters after Freshman year are still the worst semesters for high status sororities but the difference relative to other periods is less pronounced. In Spring semesters after Freshman year, the partial effect of joining a high status sorority is -.125 grade points or 18.4% of a standard deviation. This indicates recruiting and training new members still places a substantial burden on high status sororities. However, the results suggest new member education for high status sororities is also quite distracting for new members. The partial effect of joining a high status sorority is -.102 grade points in the Freshman Spring semester. Interestingly, although there is strong positive selection into high status sororities it is slightly weaker than the positive selection into low status sororities.

These heterogeneous results by social status also have important implications for policy: First, when regulating new member education, administrators should give extra scrutiny to organizations with higher social status. These organizations appear more willing to subject their recruits to new member education that distracts from coursework.

8 Conclusion

There is a heated debate about whether Greek organizations contribute to or detract from university missions. However, limited evidence exists on the effects of Greek organizations on one principal aspect of university missions: academic development of students. Measuring these effects is challenging because preferences, personality traits, and other unobserved factors which influence decisions to join Greek organizations may also affect academic outcomes. Treatment must be separated from selection to avoid spurious conclusions.

We exploit the uncommon policy of deferred affiliation at Duke University to address these selection concerns. This policy prevents Greek affiliation from occurring until the Spring semester of Freshman year which ensures students earn one semester of baseline grades before joining a fraternity or sorority. Intuitively, this allows us to employ a difference-in-differences (DinD) approach which compares changes in academic performance for Greek students before and after affiliation to changes in performance over the same period for independent students.

One concern with using a standard DinD estimator in this context is Greek affiliation may change course choice behavior. In particular, if Greek affiliation is a distraction from coursework students may respond by choosing leniently graded courses to preserve their GPA. To address this, we estimate a total treatment effect—which allows Greek affiliation to influence course choice behavior—and a partial treatment effect—which uses course fixed effects to control for changes in grading leniency. Additionally, we show the difference between these two effects identifies the effect of Greek affiliation on grading leniency. We also extend the standard DinD estimator to accommodate censoring in the dependent variable. In our sample, 33.4% of grades are worth the maximum 4.0 grade points. This represents substantial right censoring which attenuates estimates if it is ignored.

Our results show Greek organizations are a significant distraction from coursework especially in Spring semesters when new members are recruited and educated. For male students, the burden falls primary on new members who are being educated: after controlling for changes in grading leniency, fraternities decrease performance by .216 grade points (31.7% of a standard deviation) in the Freshman Spring semester. Fraternity members mitigate these negative effects by 22.2% by choosing more lenient courses. For female students, the burden falls on senior members who are in charge of recruiting and educating new members: after controlling for changes in grading leniency, sororities decrease performance by .146 grade points (21.5% of a standard deviation) in Spring semesters after Freshman year. These effects are reduced by 25.3% by choosing more lenient courses.

To analyze the effects of Greek affiliation on different students, we estimate heterogeneous treatment effects based on whether a student's SAT scores are above or below the median. Overall, low SAT students are substantially more affected than their high SAT peers. After controlling for changes in grading leniency, fraternities decrease the performance of low SAT students by .336 grade points (49.4% of a standard deviation) in the Freshman Spring semester. In the same semester, fraternities decrease the performance of high SAT students by only .120 grade points (17.6% of a standard deviation). Similarly, in Spring semesters after Freshman year, sororities decrease the performance of low SAT students by .178 grade points (26.2% of a standard deviation) but have a statistically insignificant effect on high SAT students. We also examine whether the social status of Greek organizations influences the effects of affiliation. We find high status organizations place more burden on new members who are being educated while low status organizations are more distracting for senior members.

These findings provide useful information on the determinants of academic outcomes in higher education by showing Greek affiliation can have sizable negative effects on performance. In addition, the results offer insights into ways administrators can reform Greek policies to reduce their impacts. First, university administrators may consider reforming recruitment and new member education to make it less distracting or moving it so it does not conflict with coursework. For fraternities and high status organizations, reforms should focus more on protecting new members; for sororities and low status organizations, reforms should also consider the distracting effects on senior members. Second, the results suggest universities should be especially mindful of the negative effects of Greek life on less prepared students. Universities could monitor the academic performance of 'at risk students' in each organization and reward (or punish) the entire organization based on the performance of these students. More drastically, they could set eligibility cutoffs to prevent at risk students from ever affiliating.

In addition to providing evidence on these treatment effects, we document the bias introduced by ignoring selection on unobserved variables, changes in grading leniency, or censoring at 4.0. There is generally positive selection into Greek organizations which often cannot be explained with robust controls. This positive selection is strongest for sororities, low SAT students, and high status organizations. Ignoring this selection can result in understating the negative effects of Greek affiliation by as much as .216 grade points. Students also respond to Greek affiliation by choosing more lenient courses. These behavioral responses are stronger for sorority members and students with high SAT scores and can result in understating negative effects by as much as .073 grade points. Finally, ignoring censoring leads to estimates which are attenuated by as much as 22.9%. Research which aims to identify treatment effects in higher education should be careful to account for the confounding effects of selection on unobserved factors, changes in course choice behavior, and censoring at 4.0.

References

"Academic Data." 2015. Duke University Student Affairs: Office of Fraternity & Sorority Life. Retrieved from: https://studentaffairs.duke.edu/greek/resources/academic-data

Aikins, R. D., 2011. "Academic performance enhancement: A qualitative study of the perceptions and habits of prescription stimulant-using college students." Journal of College Student Development 52, 560-576.

Ameniya, T., 1984. "Tobit Models: A Survey." Journal of Econometrics 24, 3-61.

Angrist, J., Imbens, G., Rubin, D., "Identification of Causal Effects Using Instrumental Variables." Journal of the American Statistical Association 91, 444-472.

Arcidiacono, P., Aucejo, E., Spenner, K., 2012. "What happens after enrollment? An analysis of the time path of racial differences in GPA and major choice." IZA Journal of Labor Economics 1(5).

Ashenfelter, O., 1978. "Estimating the effect of training programs on earnings." The Review of Economics and Statistics 60(1), 47-57.

Ashenfelter, O., Card, D., 1985. "Using the Longitudinal Structure of Earnings to Estimate the Effects of Training Programs." Review of Economics and Statistics 67, 648–660.

Astin, A. W., 1993. "What matters in college? Four critical years revisited." San Francisco: Jossey-Bass.

Austin, P., Escobar, M., Kopec, J., 2000. "The use of the Tobit model for analyzing measures of health status." Quality of Life Research 9, 901-910.

Bertrand, M., Duflo, E., Mullainathan, S., 2004. "How much should we trust Difference-in-Differences estimates?" The Quarterly Journal of Economics 119(1), 249-275.

Blundell, R., Costa Dias, M., 2009. "Alternative approaches to evaluation in empirical microeconomics." Journal of Human Resources 44(3), 565-640.

Bound, J., Lovenheim, M., Turner, S., 2010. "Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources." American Economic Journal: Applied Economics 2, 129-157.

Bound, J., Turner, S., 2007. “Cohort crowding: How resources affect collegiate attainment.” Journal of Public Economics 91, 877-899.

Bryant, A., Spenner, K., Martin, N., 2006. “The Campus Life and Learning Project: A Report on the First Two College Years.” Duke University.

Bryant, A., Spenner, K., Martin, N., 2007. “The Campus Life and Learning Project: A Report on the College Career.” Duke University.

Chay, K., Powell, J., 2001. “Semiparametric Censored Regression Models.” Journal of Economic Perspectives 15(4), 29-42.

“Course Selection.” 2015. Duke University Academic Advising Center. Retrieved from http://advising.duke.edu/course-selection

DeBonis, M., Shapiro, T. R., 2014, November 22. “U-Va. president suspends fraternities until Jan. 9 in wake of rape allegations.” The Washington Post.

Duflo, E., 2001. “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment.” The American Economic Review 91(4), 795-813.

Duflo, E., Glennerster, R., Kremer, M., 2007. “Using randomization in development economics research: A toolkit.” In Handbook of Development Economics Vol 4, ed. Schultz, T. P., Strauss, J., 3895-3962.

“Fraternity & Sorority Life.” 2015. Duke University Student Affairs. Retrieved from https://studentaffairs.duke.edu/greek

Grubb, F., 2006. “Does going Greek impair undergraduate academic performance?” American Journal of Economics and Sociology 65(5), 1085-1110.

Hayashi, F., 2000. “Econometrics.” Princeton University Press. Princeton, NJ.

Hayek, J. C., Carini, R. M., O’Day, P. T., Kuh, G. D., 2002. “Triumph or tragedy: Comparing student engagement levels of members of Greek-letter organizations and other Students.” Journal of College Student Development 43, 643-663.

Jacobs, P., 2014, November 4. “Dartmouth Faculty Overwhelmingly Vote to Recommend End of College's Greek System." Business Insider.

Johnson, A., 2014, November 13. "West Virginia University Student Nolan Burch Dies from Injuries." NBC News.

Johnson, V., 2003. "Grade Inflation: A Crisis in College Education." Springer-Verlag.

Kittle, P., 2011. "A Review of Deferred Recruitment." Auburn University Division of Student Affairs.

Light, A., Strayer, W., 2000. "Determinants of College Completion: School Quality or Student Ability?" The Journal of Human Resources 35(2), 299-332.

Mansfield, H., 2001. "Grade Inflation: It's Time to Face the Facts." The Chronicle Review.

McCabe, D. L., Trevino, L. K., 1997. "Individual and contextual influences on academic dishonesty: A multicampus investigation." Research in Higher Education 38, 379-396.

"Most Students in Fraternities." 2013. US News. Retrieved from http://colleges.usnews.rankingsandreviews.com/best-colleges/rankings/most-frats

"Most Students in Sororities." 2013. US News. Retrieved from http://colleges.usnews.rankingsandreviews.com/best-colleges/rankings/most-sororities

Pascarella, E. T., Terenzini, P. T., 2005. "How college affects students: Volume 2, a third decade of research." San Francisco: Jossey-Bass.

Passow, H. J., Mayhew, M. J., Finelli, C. J., Harding, T. S., Carpenter, D. D., 2006. "Factors influencing engineering students' decisions to cheat by type of assessment." Research in Higher Education 47, 643-684.

Perez-Pena, R., Yaccino, S., 2014, June 12. "Efforts by Colleges to Curb Assaults Focus on Fraternities." New York Times.

Pike, G. R., 2000. "The influence of fraternity or sorority membership on students' college experiences and cognitive development." Research in Higher Education 41, 117-139.

Pike, G. R., 2003. "Membership in a fraternity or sorority, student engagement, and educational outcomes at AAU public research universities." Journal of College Student Development 44, 369-382.

Pike, G. R., Askew, J., 1990. “The impact of fraternity or sorority membership on academic involvement and learning outcomes.” NAPSA Journal 28, 13-19.

Puhani, P., 2012. “The treatment effect, the cross difference, and the interaction term in non-linear ’difference-in-differences’ models.” Economics Letters 115, 85-87.

Robinson, P., 1982. “On the Asymptotic Properties of Estimators of Models Containing Limited Dependent Variables.” Econometrica 50(1), 27-41.

Rockey, D. L., Beason, K. R., Howington, E. B., Rockey, C. M., 2005. “Gambling by Greek-affiliated college students: An association between affiliation and gambling.” Journal of College Student Development 46, 75-87.

Sabot, R., Wakemann-Linn, J., 1991. “Grade inflation and course choice.” Journal of Economic Perspectives 5(1), 159-170.

Stinebrickner, R., Stinebrickner, T., 2003. “Working during School and Academic Performance.” Journal of Labor Economics 21(2), 473-491.

Stinebrickner, R., Stinebrickner, T., 2004. “Time-use and college outcomes.” Journal of Econometrics 121(1-2), 243-269.

Thompson, J. G., Oberle, C. D., Liley, J. L., 2011. “Self-efficacy and learning in sorority and fraternity students.” Journal of College Student Development 52, 749-753.

Tobin, J., 1958. “Estimation of Relationships for Limited Dependent Variables.” Econometrica 26(1), 24-36.

Todd, P., Wolpin, K., 2003. “On the specification and estimation of the production function for cognitive achievement.” The Economic Journal 113, F3-F33.

Tuchler, M., 2012, February 9. “Hazing incidents increase.” Duke Chronicle.

Walker, J., Martin, N., Hussey, A., 2015. “Greek Organization Membership and Collegiate Oucomes at an Elite, Private University.” Research in Higher Education 56, 203-227.

Wechsler, H., Kuh, G. D., Davenport, A. E., 1996. “Fraternities, sororities, and binge drinking: Results from a national study of American colleges.” NAPSA Journal 33, 260-279.

Williams, A. E., Janosik, S. M., 2007. “An examination of academic dishonesty among sorority and nonsorority women." Journal of College Student Development 48, 706-714.

Zwick, R., 2004. "Is the SAT a 'Wealth Test?' The Link between Educational Achievement and Socioeconomic Status." In Rethinking the SAT: The Future of Standardized Testing in University Admissions, ed. R. Zwick. RoutledgeFalmer. New York, NY.

Table 1: Greek and Independent Populations by Gender

	Female			**Male**		
	Independent	Greek	P-Value	Independent	Greek	P-Value
Race						
White	46.8%	71.3%	0.000	61.2%	72.6%	0.011
Black	20.8%	6.3%	0.000	7.1%	3.9%	0.050
Hispanic	7.2%	7.9%	0.693	7.1%	8.7%	0.396
Asian	17.9%	5.3%	0.000	19.6%	9.1%	0.001
Other Race	7.3%	9.3%	0.496	5.0%	5.7%	0.794
Admissions						
SAT Composite Score	1,373	1,384	0.367	1,420	1,440	0.064
Achievement	4.25	4.20	0.534	4.22	4.21	0.895
Curriculum	4.60	4.63	0.666	4.69	4.74	0.400
Essay	3.38	3.42	0.447	3.38	3.40	0.692
Personal Qualities	3.44	3.45	0.888	3.37	3.42	0.456
Recommendation	3.73	3.75	0.820	3.73	3.75	0.703
Demographics						
Private HS	15.4%	25.8%	0.015	19.8%	21.2%	0.747
Legacy Student	16.0%	24.3%	0.060	17.7%	14.7%	0.462
Less than Bachelor's	13.6%	4.9%	0.000	8.4%	6.6%	0.473
Bachelor's or Masters	49.7%	47.9%	0.720	50.0%	45.9%	0.447
Professional	34.0%	42.2%	0.102	37.1%	44.5%	0.177
$0 - $49,999	18.7%	10.1%	0.006	14.5%	7.7%	0.015
$50,000 - $99,999	20.8%	14.3%	0.073	21.4%	11.0%	0.004
$100,000 - $199,999	27.8%	24.2%	0.433	30.5%	30.7%	0.969
$200,000+	21.7%	34.5%	0.007	23.8%	39.9%	0.002
Performance						
Fresh Fall GPA	3.240	3.367	0.004	3.290	3.359	0.199
Fresh Spring GPA	3.256	3.336	0.089	3.275	3.220	0.322

Notes: P-values report the probabilities that Greek and independent population means are equal. Achievement, Curriculum, Essay, Personal Qualities, and Recommendation are internal ratings of application quality; possible values are 1 through 5. Education measures are for the student's most educated parent; omitted category is missing data. Income measures are for the student's household.

Table 2: The Effects of Greek Affiliation on Academic Outcomes (Censored Regression)

	No Controls			Demographics			add Baseline Prep		
	Total	Partial	Leniency	Total	Partial	Leniency	Total	Partial	Leniency
Fraternity Member									
Fresh Spring	-0.158***	-0.22***	0.062**	-0.169***	-0.222***	0.053*	-0.168***	-0.216***	0.048
	0.0549	*0.0511*	*0.0303*	*0.0578*	*0.0532*	*0.0314*	*0.0578*	*0.0523*	*0.0312*
Fall after Fresh yr	-0.0759	-0.118**	0.0421	-0.0983	-0.137**	0.0387	-0.0993*	-0.132**	0.0327
	0.0578	*0.0531*	*0.0268*	*0.0598*	*0.0545*	*0.0277*	*0.0599*	*0.0537*	*0.0277*
Spring after Fresh yr	-0.0582	-0.0851	0.0269	-0.0675	-0.0921	0.0246	-0.0688	-0.089	0.0202
	0.0600	*0.0555*	*0.0260*	*0.0621*	*0.0566*	*0.0263*	*0.0623*	*0.0562*	*0.0264*
Selection	0.0679	0.102	-0.0341	0.0643	0.098	-0.0337	0.0373	0.0627	-0.0254
	0.0743	*0.0722*	*0.0229*	*0.0727*	*0.0703*	*0.0233*	*0.0649*	*0.0620*	*0.0231*
Sorority Member									
Fresh Spring	-0.0557	-0.0889**	0.0332	-0.0578	-0.0874*	0.0296	-0.0578	-0.0818*	0.024
	0.0456	*0.0446*	*0.0281*	*0.0466*	*0.0455*	*0.0291*	*0.0468*	*0.0462*	*0.0284*
Fall after Fresh yr	-0.0238	-0.0854*	0.0616***	-0.037	-0.101**	0.064***	-0.0336	-0.0947**	0.0611**
	0.0466	*0.0437*	*0.0236*	*0.0503*	*0.0449*	*0.0245*	*0.0506*	*0.0454*	*0.0244*
Spring after Fresh yr	-0.106**	-0.138***	0.032	-0.111**	-0.146***	0.035	-0.109**	-0.146***	0.037
	0.0453	*0.0419*	*0.0248*	*0.0471*	*0.0421*	*0.0253*	*0.0472*	*0.0417*	*0.0253*
Selection	0.128**	0.166***	-0.038*	0.0743	0.116**	-0.0417*	0.0828	0.122**	-0.0392*
	0.0542	*0.0516*	*0.0222*	*0.0558*	*0.0531*	*0.0226*	*0.0510*	*0.0495*	*0.0223*
Controls									
Demographics (time varying effects)	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Preparation	No	No	No	No	No	No	Yes	Yes	Yes
Student Observations	1,008	1,008	1,008	1,008	1,008	1,008	1,008	1,008	1,008
Course Observations	29,001	29,001	29,001	29,001	29,001	29,001	29,001	29,001	29,001

Notes: Block bootstrap standard errors (300 iterations) in italics. *** p<0.01, ** p<0.05, * p<0.1. Total effect allows for changes in course choice behavior; Partial effect controls for changes in course choice behavior; Leniency effect measures effects on grading leniency. Demographic controls include indicator variables for education of the most educated parent, family income, student race, whether the student is a legacy, and whether the student attended a private high school. Baseline preparation controls include composite SAT scores and five internal measures of application quality.

Table 3: The Effect of Greek Affiliation on Grades (OLS)

	No Controls			Demographics			add Baseline Prep		
	Total	Partial	Leniency	Total	Partial	Leniency	Total	Partial	Leniency
Fraternity Member									
Fresh Spring	-0.124***	-0.179***	0.055**	-0.131***	-0.179***	0.048*	-0.131***	-0.174***	0.043*
	0.0437	*0.0408*	*0.0236*	*0.0465*	*0.0428*	*0.0246*	*0.0465*	*0.0422*	*0.0245*
Fall after Fresh yr	-0.0585	-0.0919**	0.0334*	-0.0761*	-0.109***	0.0329	-0.0761*	-0.105***	0.0289
	0.0426	*0.0392*	*0.0201*	*0.0443*	*0.0405*	*0.0209*	*0.0443*	*0.0400*	*0.0207*
Spring after Fresh yr	-0.051	-0.0754*	0.0244	-0.0591	-0.0814*	0.0223	-0.0591	-0.0774*	0.0183
	0.0450	*0.0419*	*0.0197*	*0.0470*	*0.0432*	*0.0200*	*0.0470*	*0.0429*	*0.0200*
Selection	0.0691	0.097*	-0.0279	0.0591	0.0891*	-0.03	0.0417	0.0659	-0.0242
	0.0560	*0.0540*	*0.0184*	*0.0552*	*0.0531*	*0.0188*	*0.0487*	*0.0464*	*0.0185*
Sorority Member									
Fresh Spring	-0.0468	-0.0732**	0.0264	-0.0479	-0.0711**	0.0232	-0.0479	-0.0682**	0.0203
	0.0346	*0.0326*	*0.0218*	*0.0354*	*0.0337*	*0.0222*	*0.0354*	*0.0340*	*0.0216*
Fall after Fresh yr	-0.0241	-0.0711**	0.047**	-0.0353	-0.0846**	0.0493***	-0.0353	-0.0861***	0.0508***
	0.0347	*0.0325*	*0.0184*	*0.0374*	*0.0335*	*0.0188*	*0.0374*	*0.0333*	*0.0184*
Spring after Fresh yr	-0.0778**	-0.109***	0.0312	-0.084**	-0.116***	0.032*	-0.084**	-0.118***	0.034*
	0.0338	*0.0314*	*0.0191*	*0.0352*	*0.0315*	*0.0193*	*0.0352*	*0.0310*	*0.0189*
Selection	0.127***	0.152***	-0.025	0.0702*	0.102**	-0.0318*	0.0767**	0.108***	-0.0313*
	0.0411	*0.0391*	*0.0179*	*0.0422*	*0.0406*	*0.0182*	*0.0382*	*0.0370*	*0.0176*
Controls									
Demographics (time varying effects)	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Preparation	No	No	No	No	No	No	Yes	Yes	Yes
Student Observations	1,008	1,008	1,008	1,008	1,008	1,008	1,008	1,008	1,008
Course Observations	29,001	29,001	29,001	29,001	29,001	29,001	29,001	29,001	29,001

Notes: Block bootstrap standard errors (300 iterations) in italics. *** p<0.01, ** p<0.05, * p<0.1. Total effect allows for changes in course choice behavior; Partial effect controls for changes in course choice behavior; Leniency effect measures effects on grading leniency. Demographic controls include indicator variables for education of the most educated parent, family income, student race, whether the student is a legacy, and whether the student attended a private high school. Baseline preparation controls include composite SAT scores and five internal measures of application quality.

Table 4: Heterogeneous Effects by Initial Preparation (Male Coefficients)

	No Controls			Demographics			add Baseline Prep		
	Total	Partial	Leniency	Total	Partial	Leniency	Total	Partial	Leniency
Low SAT Fraternity Member									
Fresh Spring	-0.334***	-0.353***	0.0190	-0.317***	-0.338***	0.0210	-0.32***	-0.336***	0.0160
	0.0792	*0.0725*	*0.0399*	*0.0818*	*0.0749*	*0.0415*	*0.0819*	*0.0733*	*0.0421*
Fall after Fresh yr	-0.21***	-0.248***	0.0380	-0.201**	-0.241***	0.0400	-0.206**	-0.245***	0.0390
	0.0801	*0.0774*	*0.0363*	*0.0813*	*0.0767*	*0.0371*	*0.0812*	*0.0751*	*0.0363*
Spring after Fresh yr	-0.139*	-0.183**	0.0440	-0.1240	-0.17**	0.0460	-0.1260	-0.164**	0.0380
	0.0775	*0.0818*	*0.0349*	*0.0783*	*0.0804*	*0.0359*	*0.0786*	*0.0798*	*0.0351*
Selection	0.247***	0.288***	-0.0410	0.199**	0.243***	-0.0440	0.175**	0.216***	-0.0410
	0.0917	*0.0922*	*0.0314*	*0.0882*	*0.0887*	*0.0320*	*0.0791*	*0.0794*	*0.0317*
High SAT Fraternity Member									
Fresh Spring	-0.0153	-0.1070	0.0917**	-0.0465	-0.125*	0.0785*	-0.0472	-0.12*	0.0728*
	0.0718	*0.0690*	*0.0396*	*0.0742*	*0.0707*	*0.0409*	*0.0741*	*0.0701*	*0.0401*
Fall after Fresh yr	0.0357	-0.0114	0.0471	-0.0106	-0.0528	0.0422	-0.0105	-0.0435	0.0330
	0.0777	*0.0723*	*0.0369*	*0.0808*	*0.0742*	*0.0379*	*0.0809*	*0.0735*	*0.0376*
Spring after Fresh yr	0.0097	-0.0079	0.0176	-0.0191	-0.0327	0.0136	-0.0202	-0.0283	0.0081
	0.0789	*0.0694*	*0.0379*	*0.0830*	*0.0732*	*0.0381*	*0.0830*	*0.0725*	*0.0380*
Selection	-0.0896	-0.0619	-0.0277	-0.0589	-0.0322	-0.0267	-0.0669	-0.0513	-0.0156
	0.0978	*0.0952*	*0.0306*	*0.0964*	*0.0937*	*0.0312*	*0.0883*	*0.0844*	*0.0308*
Controls									
Demographics (time varying effects)	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Preparation	No	No	No	No	No	No	Yes	Yes	Yes
Student Observations	1,008	1,008	1,008	1,008	1,008	1,008	1,008	1,008	1,008
Course Observations	29,001	29,001	29,001	29,001	29,001	29,001	29,001	29,001	29,001

Notes: Block bootstrap standard errors (300 iterations) in italics. *** p<0.01, ** p<0.05, * p<0.1. Total effect allows for changes in course choice behavior; Partial effect controls for changes in course choice behavior; Leniency effect measures effects on grading leniency. Demographic controls include indicator variables for education of the most educated parent, family income, student race, whether the student is a legacy, and whether the student attended a private high school. Baseline preparation controls include composite SAT scores and five internal measures of application quality.

Table 5: Heterogeneous Effects by Initial Preparation (Female Coefficients)

	No Controls			Demographics			add Baseline Prep		
	Total	Partial	Leniency	Total	Partial	Leniency	Total	Partial	Leniency
Low SAT Sorority Member									
Fresh Spring	-0.0564	-0.0694	0.0130	-0.0459	-0.0651	0.0192	-0.0470	-0.0642	0.0172
	0.0571	*0.0548*	*0.0296*	*0.0565*	*0.0557*	*0.0312*	*0.0565*	*0.0559*	*0.0312*
Fall after Fresh yr	-0.0730	-0.137**	0.064**	-0.0759	-0.147***	0.0711**	-0.0730	-0.144***	0.071**
	0.0585	*0.0550*	*0.0281*	*0.0590*	*0.0550*	*0.0293*	*0.0591*	*0.0553*	*0.0295*
Spring after Fresh yr	-0.145**	-0.168***	0.0230	-0.145**	-0.178***	0.0330	-0.143**	-0.178***	0.0350
	0.0619	*0.0574*	*0.0281*	*0.0610*	*0.0574*	*0.0294*	*0.0610*	*0.0569*	*0.0293*
Selection	0.221***	0.263***	-0.0420	0.146**	0.196***	-0.05*	0.137**	0.186***	-0.049*
	0.0657	*0.0608*	*0.0269*	*0.0630*	*0.0591*	*0.0278*	*0.0594*	*0.0565*	*0.0275*
High SAT Sorority Member									
Fresh Spring	-0.0504	-0.1080	0.0576	-0.0797	-0.1160	0.0363	-0.0778	-0.1100	0.0322
	0.0806	*0.0830*	*0.0479*	*0.0829*	*0.0845*	*0.0486*	*0.0826*	*0.0837*	*0.0477*
Fall after Fresh yr	0.0673	0.0129	0.0544	0.0254	-0.0140	0.0394	0.0265	-0.0120	0.0385
	0.0895	*0.0841*	*0.0402*	*0.0949*	*0.0878*	*0.0416*	*0.0952*	*0.0882*	*0.0422*
Spring after Fresh yr	-0.0325	-0.0850	0.0525	-0.0642	-0.0999	0.0357	-0.0628	-0.0990	0.0362
	0.0739	*0.0758*	*0.0416*	*0.0782*	*0.0789*	*0.0424*	*0.0780*	*0.0782*	*0.0431*
Selection	-0.0426	-0.0174	-0.0252	-0.0099	0.0072	-0.0171	0.0049	0.0220	-0.0171
	0.0968	*0.0963*	*0.0354*	*0.0992*	*0.0998*	*0.0359*	*0.0948*	*0.0945*	*0.0364*
Controls									
Demographics	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
(time varying effects)									
Baseline Preparation	No	No	No	No	No	No	Yes	Yes	Yes
Student Observations	1,008	1,008	1,008	1,008	1,008	1,008	1,008	1,008	1,008
Course Observations	29,001	29,001	29,001	29,001	29,001	29,001	29,001	29,001	29,001

Notes: Block bootstrap standard errors (300 iterations) in italics. *** p<0.01, ** p<0.05, * p<0.1. Total effect allows for changes in course choice behavior; Partial effect controls for changes in course choice behavior; Leniency effect measures effects on grading leniency. Demographic controls include indicator variables for education of the most educated parent, family income, student race, whether the student is a legacy, and whether the student attended a private high school. Baseline preparation controls include composite SAT scores and five internal measures of application quality.

Table 6: Heterogeneous Effects by Social Status (Male Coefficients)

	No Controls			Demographics			add Baseline Prep		
	Total	Partial	Leniency	Total	Partial	Leniency	Total	Partial	Leniency
Low Status Fraternity Member									
Fresh Spring	-0.136*	-0.176**	0.0400	-0.142*	-0.18**	0.0380	-0.144*	-0.172**	0.0280
	0.0698	*0.0699*	*0.0390*	*0.0738*	*0.0720*	*0.0410*	*0.0736*	*0.0706*	*0.0409*
Fall after Fresh yr	-0.1160	-0.16**	0.0440	-0.131*	-0.172**	0.0410	-0.132*	-0.16**	0.0280
	0.0757	*0.0762*	*0.0319*	*0.0795*	*0.0778*	*0.0322*	*0.0796*	*0.0759*	*0.0324*
Spring after Fresh yr	-0.0917	-0.1120	0.0203	-0.0957	-0.1140	0.0183	-0.0976	-0.1060	0.0084
	0.0831	*0.0879*	*0.0320*	*0.0845*	*0.0875*	*0.0316*	*0.0844*	*0.0857*	*0.0316*
Selection	0.0108	0.0291	-0.0183	0.0199	0.0412	-0.0213	-0.0367	-0.0258	-0.0109
	0.0923	*0.0964*	*0.0301*	*0.0930*	*0.0957*	*0.0298*	*0.0789*	*0.0793*	*0.0294*
High Status Fraternity Member									
Fresh Spring	-0.175***	-0.257***	0.082**	-0.191***	-0.258***	0.067*	-0.19***	-0.254***	0.0640
	0.0661	*0.0601*	*0.0391*	*0.0674*	*0.0604*	*0.0400*	*0.0677*	*0.0596*	*0.0393*
Fall after Fresh yr	-0.0439	-0.0834	0.0395	-0.0703	-0.106*	0.0357	-0.0722	-0.109*	0.0368
	0.0710	*0.0623*	*0.0358*	*0.0718*	*0.0627*	*0.0371*	*0.0719*	*0.0624*	*0.0361*
Spring after Fresh yr	-0.031	-0.0632	0.0322	-0.0437	-0.0726	0.0289	-0.0443	-0.074	0.0297
	0.0677	*0.0553*	*0.0362*	*0.0695*	*0.0560*	*0.0375*	*0.0699*	*0.0561*	*0.0369*
Selection	0.1140	0.16*	-0.0460	0.1010	0.144*	-0.0430	0.0989	0.136*	-0.0371
	0.0918	*0.0885*	*0.0298*	*0.0874*	*0.0834*	*0.0305*	*0.0806*	*0.0764*	*0.0291*
Controls									
Demographics (time varying effects)	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Preparation	No	No	No	No	No	No	Yes	Yes	Yes
Student Observations	1,008	1,008	1,008	1,008	1,008	1,008	1,008	1,008	1,008
Course Observations	29,001	29,001	29,001	29,001	29,001	29,001	29,001	29,001	29,001

Notes: Block bootstrap standard errors (300 iterations) in italics. *** p<0.01, ** p<0.05, * p<0.1. Total effect allows for changes in course choice behavior; Partial effect controls for changes in course choice behavior; Leniency effect measures effects on grading leniency. Demographic controls include indicator variables for education of the most educated parent, family income, student race, whether the student is a legacy, and whether the student attended a private high school. Baseline preparation controls include composite SAT scores and five internal measures of application quality.

Table 7: Heterogeneous Effects by Social Status (Female Coefficients)

	No Controls			Demographics			add Baseline Prep		
	Total	Partial	Leniency	Total	Partial	Leniency	Total	Partial	Leniency
Low Status Sorority Member									
Fresh Spring	0.0005	-0.0616	0.0621	-0.0064	-0.0639	0.0575	-0.0058	-0.0522	0.0464
	0.0679	*0.0620*	*0.0387*	*0.0683*	*0.0633*	*0.0380*	*0.0685*	*0.0648*	*0.0365*
Fall after Fresh yr	-0.0538	-0.127*	0.0732**	-0.0714	-0.14**	0.0686**	-0.0682	-0.128*	0.0598*
	0.0733	*0.0713*	*0.0316*	*0.0746*	*0.0698*	*0.0316*	*0.0751*	*0.0698*	*0.0308*
Spring after Fresh yr	-0.111*	-0.175***	0.064**	-0.12*	-0.182***	0.062*	-0.119*	-0.176***	0.057*
	0.0636	*0.0621*	*0.0321*	*0.0614*	*0.0577*	*0.0329*	*0.0614*	*0.0564*	*0.0327*
Selection	0.1310	0.187**	-0.056*	0.131*	0.187***	-0.056*	0.0892	0.142**	-0.0528*
	0.0801	*0.0756*	*0.0300*	*0.0740*	*0.0697*	*0.0293*	*0.0715*	*0.0689*	*0.0284*
High Status Sorority Member									
Fresh Spring	-0.0904*	-0.105*	0.0146	-0.0928*	-0.102*	0.0092	-0.0934*	-0.102*	0.0086
	0.0537	*0.0552*	*0.0328*	*0.0548*	*0.0548*	*0.0342*	*0.0551*	*0.0549*	*0.0335*
Fall after Fresh yr	-0.0049	-0.0594	0.05448*	-0.0124	-0.0732	0.0608**	-0.0084	-0.0716	0.06321**
	0.0521	*0.0493*	*0.0286*	*0.0559*	*0.0508*	*0.0304*	*0.0562*	*0.0517*	*0.0299*
Spring after Fresh yr	-0.102*	-0.115**	0.0130	-0.103*	-0.12**	0.0170	-0.101*	-0.125**	0.0240
	0.0548	*0.0517*	*0.0290*	*0.0573*	*0.0526*	*0.0296*	*0.0575*	*0.0524*	*0.0289*
Selection	0.127**	0.153***	-0.0260	0.0382	0.0696	-0.0314	0.0820	0.113**	-0.0310
	0.0608	*0.0587*	*0.0274*	*0.0648*	*0.0623*	*0.0282*	*0.0570*	*0.0554*	*0.0272*
Controls									
Demographics (time varying effects)	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Preparation	No	No	No	No	No	No	Yes	Yes	Yes
Student Observations	1,008	1,008	1,008	1,008	1,008	1,008	1,008	1,008	1,008
Course Observations	29,001	29,001	29,001	29,001	29,001	29,001	29,001	29,001	29,001

Notes: Block bootstrap standard errors (300 iterations) in italics. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Total effect allows for changes in course choice behavior; Partial effect controls for changes in course choice behavior; Leniency effect measures effects on grading leniency. Demographic controls include indicator variables for education of the most educated parent, family income, student race, whether the student is a legacy, and whether the student attended a private high school. Baseline preparation controls include composite SAT scores and five internal measures of application quality.

Table 8: Full Sample Composition (ONLINE APPENDIX)

	Female				Male			
	Independent	P-Value	Greek	P-Value	Independent	P-Value	Greek	P-Value
Race								
White	46.7%	0.869	72.2%	0.827	61.3%	0.941	71.8%	0.938
Black	20.7%	0.932	5.7%	0.765	7.3%	0.896	4.1%	0.898
Hispanic	6.8%	0.703	7.6%	0.865	7.2%	0.893	9.8%	0.700
Asian	18.1%	0.910	5.8%	0.832	19.3%	0.966	9.0%	0.914
Other Race	7.8%	0.785	8.7%	0.860	5.0%	0.989	5.2%	0.865
Admissions								
SAT Composite Score	1,374	0.953	1,386	0.912	1,418	0.816	1,433	0.554
Achievement	4.25	0.978	4.16	0.797	4.21	0.771	4.17	0.749
Curriculum	4.60	0.953	4.65	0.692	4.68	0.739	4.72	0.733
Essay	3.39	0.993	3.40	0.801	3.38	0.874	3.38	0.698
Personal Qualities	3.44	0.889	3.44	0.889	3.37	0.954	3.41	0.956
Recommendation	3.73	0.924	3.74	0.981	3.73	0.957	3.73	0.772
Demographics								
Private HS	15.9%	0.978	26.6%	0.940	20.8%	0.802	21.5%	0.992
Legacy Student	15.9%	0.921	25.5%	0.786	18.8%	0.699	13.2%	0.763
Less than Bachelor's	13.3%	0.945	5.5%	0.790	8.5%	0.972	5.9%	0.815
Bachelor's or Masters	48.4%	0.738	46.6%	0.844	50.4%	0.931	47.3%	0.755
Professional	35.2%	0.767	43.1%	0.913	37.1%	0.987	43.6%	0.874
$0 - $49,999	18.3%	0.962	10.8%	0.792	13.9%	0.870	7.5%	0.937
$50,000 - $99,999	21.4%	0.851	13.9%	0.941	22.0%	0.980	11.0%	0.954
$100,000 - $199,999	27.5%	0.877	23.0%	0.812	30.7%	0.964	31.2%	0.914
$200,000+	21.5%	0.919	35.2%	0.952	24.2%	0.863	39.7%	0.966

Notes: P-values report the probabilities that full sample and estimation sample means are equal. Achievement, Curriculum, Essay, Personal Qualities, and Recommendation are internal ratings of application quality; possible values are 1 through 5. Education measures are for the student's most educated parent; omitted category is missing data. Income measures are for the student's household.