Do Differential Grading Norms Across Fields Matter for Major Choice? Evidence from a Policy Change in Florida

Veronica Minaya

May 6, 2017

Abstract

Using administrative data from the Florida Department of Education, this paper examines the effect of changing the grading scale from whole-letter grades to plus/minus grades on STEM major choice. I rely on a difference-in-differences framework that compares before and after a grade policy change at two institutions to similar students at other institutions over the same time period. I find that an arbitrary change in the grading scale significantly reduces grading differentials and increases the likelihood of students graduating with a STEM degree. These results represent the first direct, quasi-experimental evidence regarding the effect of a grade scale change on STEM major choice.

Veronica Minaya is a visiting scholar at Teachers College at Columbia University. The author gratefully acknowledges the Florida Department of Education for providing access to the restricted data utilized herein. The author thanks Judith Scott-Clayton, Peter Bergman, Randall Reback and Jonah Rockoff for their feedback on early drafts. The opinions expressed are those of the author and do not represent views Florida Department of Education. The author cannot share the data used herein due to confidentiality agreements.

1. Introduction

Gender and racial disparities in STEM graduation and major choice are stark and studies have frequently found that women and racial minorities (hereafter called STEM minorities) leave STEM fields at higher rates than their counterparts (Anderson & Kim, 2006; Hill et al., 2010; Griffith, 2010; Huang, Taddese, & Walter, 2000; Kokkelenberg & Sinha, 2010; Shaw & Barbuti, 2010).[1] Since STEM degrees pay substantially and increasingly more than other fields (Altonji et al., 2012; Altonji et al., 2014), the high attrition rates among STEM minorities are believed to be a driving factor behind the gender and race wage gap (Gerber & Cheung, 2008; Brown & Corcoran, 1996; Thomas, 1985).

In this paper, I evaluate the effect of a change in grading scale on patterns of college grades in lower-division courses and STEM graduation and major choice. In the early 2000s, two universities in Florida, the University of Central Florida (UCF) and University of South Florida (USF), changed their grading scale from whole-letter to plus/minus grades. Previous research on changing the grading scale has been mostly descriptive and produced in the context of curbing grade inflation. I contribute to the literature by providing the first quasi-experimental estimates of the effect of a change in the grading scale rather than on other academic and more policy-relevant outcomes such as major choices, persistence, and graduation rates. I also improve on this prior work by identifying students who take STEM and non-STEM courses and analyzing if

[1] The persistence rate of women in STEM is less than that of men, and proportionally more females than males left STEM fields by mostly switching to a non-STEM major (32 percent vs. 26 percent). It is worth noting that proportionally more males than females left STEM fields by dropping out of college (24 percent vs. 14 percent) (Chen, 2013). Similarly, the gap in STEM attrition rates between racial minorities and non-minorities is alarming. One-third of White students and 42 percent of Asian-American students who started college as intended STEM majors graduated with STEM degrees by the end of five years, compared to 22 percent of Latino students, 18 percent of Black students, and 20 percent of Native American (Hurtado et al., 2010).

the effect of this grading policy is differentiated across STEM and non-STEM departments (low- and high-grading departments) and lower-division courses.

I specifically test three related hypotheses. First, all else equal, a more continuous and refined grading scale will lead to reduce grade differentials between STEM and non-STEM courses. I find that a change in the grading scale from whole-grade grading scale to plus/minus grading has a substantial effect on first year grade differentials between STEM and non-STEM lower-division courses. Although not obvious a priori, this policy had the effect of substantially increasing grades in STEM fields. Students who attended institutions that changed their grading scale experienced a smaller difference between their STEM and non-STEM GPAs during the first year of enrollment than similar students attending institutions that did not implement any chance in their grading scale. I use these results to introduce my second hypothesis and to frame a discussion of the mechanisms potentially responsible for reducing STEM attrition.

Second, a reduction in grading disparities will make students more attracted to STEM, which in turn will improve STEM graduation and major choice. I find that after the grading policy change, students attending treated institutions are significantly more likely to graduate in STEM in 6 years or less, or to choose a STEM major at the beginning of their second and third year of enrollment. One possible reason to explain these results is that students might be vulnerable to grades differentials due to differences in the cost of study and grading standards between STEM and non-STEM fields (Arcidiacono, 2011; Barnes, Bull, Campbell, & Perry, 2001; K. Rask, 2010; Johnson, 2003).

Finally, a reduction in grading disparities will lead to an improvement in the gender and race gap in STEM attrition, based on the empirical evidence that suggests that STEM minorities value grades more than their counterparts (Rask & Bailey, 2002; Rask & Tiefenthaler,2008). I find significant impacts by gender and race, with a similar pattern of positive effects on STEM graduation and major choice outcomes for all subgroups, which is consistent with the fact that grade differentials for all subgroups were equally reduced after the change in the grading scale.

The importance of grades and the distortions in major choice decisions have been explored by researchers. Grades reinforce or alter students' initial expectations and preferences, which, in turn, influence their final course and major choices (Attewell, Heil, & Reisel, 2011; Crisp et al., 2009; Stinebrickner & Stinebrickner, 2011). In particular, if students enter college with incorrect information regarding their relative strengths, grading provides a potentially effective mechanism for informing a student which major they should choose (Stinebrickner & Stinebrickner, 2009). Lower grades have been linked to major attrition in general (DeBerard, Spielmans, & Julka, 2004; Nora, Cabrera, Hagedorn, & Pascarella, 1996; Reason, 2009) and in STEM specifically (Strenta, Elliott, Adair, Matier, & Scott, 1994; Ost, 2010; Crisp et al., 2009; K. Rask, 2010).

Disparities in grading policies across departments have motivated a scant literature exploring the effects of equalizing grades on course and major choices. The starting point of this literature is the observation that STEM departments are associated with higher and more rigorous grading standards than non-STEM departments (Arcidiacono, 2011; Barnes, Bull, Campbell, & Perry, 2001; K. Rask, 2010; Johnson, 2003). This emerging grading gap between STEM and non-STEM departments could be partially explained by grade inflation and its disproportionate effect on non-STEM grades (Ost, 2010). This suggests that concurrent with a potential disconnect between grades and ability, there is an argument to be made that grade differentials between

STEM and non-STEM courses might be influencing STEM major choice by exacerbating the importance of relative grades as a measure of relative course performance.[2]

Equalizing average grades differentials between STEM and non-STEM courses may potentially improve the number and composition of STEM graduates (Sabot & Wakeman-Linn, 1991; K. Rask, 2010; Arcidiacono et al., 2015b). A discrepancy between grades in STEM and non-STEM courses may discourage students from majoring in subjects from low-grading departments or STEM fields,[3] by pushing them away from their STEM intended major when obtaining high grades in non-STEM courses (Ost, 2010). Since STEM minorities respond more strongly to grade incentives than do non-minorities (Rask & Bailey, 2002; Rask & Tiefenthaler, 2008), reducing disparities in grading policies might not only improve STEM graduation, but also the gender and race gap in STEM attrition.

Concerns over grading disparities and grade inflation have led institutions to adopt various grading policies and researchers to investigate their effects. There is evidence that grading policies affect course and major choice by reducing grade inflation. Butcher et al. (2014) evaluate the consequences of Wellesley's policy, which was intended to cap the fraction of A's, by comparing outcomes in high-grading departments that were obligated to lower their grades with outcomes in low-grading departments that were not. The authors find that the policy brings average grades down in high-grading departments, reducing compression at the top of the grade

[2] For a given student relative grades measure relative performance in introductory STEM courses relative to the same student in other non-STEM courses. The literature has studied the importance of relative performance of college students in introductory courses as a determinant of undergraduate major choice (Sabot & Wakeman-Linn, 1991; Dynan & Rouse, 1997; Robb & Robb, 1999; Chizmar, 2000; Jensen & Owen, 2001; K. Rask, 2010; K. N. Rask & Bailey, 2002; Ost, 2010; Seymour & Hewitt, 1997; Stinebrickner & Stinebrickner, 2011; K. Rask & Tiefenthaler, 2008).

[3] Rojstaczer & Healy (2010) looked at contemporary grades from over 160 colleges and universities in the United States with a combined enrollment of over 2,000,000 students and historical grades from over 80 schools. They conclude that nationally for all colleges and universities, science departments grade on average roughly 0.4 lower on a 4.0 scale than humanities departments and 0.2 lower than social science departments.

distribution. They also show that this policy shifts towards science classes and science majors, indicating that students' choices about their majors are linked with grades.

Research on changing the grading scale have yielded conflicting results. While some authors found a reduction in grade inflation using plus/minus grading (Shannon, 1979; Farland & Cepeda, 1989; Bressette, 2002), others have reported no effect (Baker III & Bates, 1999). This research, however, also points out that the effect of this grading policy could be differentiated across courses and departments (Bressette, 2002 and K. D. Barnes & Buring, 2012).

This paper aims to advance research by providing the first direct examination of the effects of changing the grading scale on STEM graduation and major choice. Using a restricted-use administrative database on Florida public college enrollees, I applied a difference-in-difference approach, comparing similar students whose grading differentials between STEM and non-STEM courses changed over time at institutions with or without the change to estimate the effect of this institutional grading policy on STEM outcomes. I have three key findings. First, I find that the impact of the change in the grading scale on the difference between STEM and non-STEM GPAs during the first year of enrollment is about 0.152, a relative increase of about a sixth of a grade differential in treated institutions. This effect is mainly explained by a substantial increase on STEM grades.

Second, after the grading policy change, students attending treated institutions were 3 percentage points more likely to graduate in STEM in 6 years or less, which corresponds to an increase of about 40 percent. These positive effects on STEM outcomes seem to be equally explained by a reduction in the probability of majoring in non-STEM and dropping out of college without earning a BA.

Finally, my analysis demonstrates that this grading policy has complex heterogeneous effects in the way students are discouraged from leaving STEM. This policy has larger effects for men than women and for racial minorities than non- minorities. However, for women the effects on STEM outcomes seem to be explained by a reduction in the probability of leaving STEM by switching into non-STEM, whereas for men these effects are paired with a larger decrease in the probability of dropping out of college. In fact, men were relatively more affected at the bottom of their grade distribution which might explain why overall impact estimates are higher for men.

Section II describes the policy background. Section III describes dataset in detail and presents descriptive statistics. Section IV presents the empirical specification. Section V presents results and subgroup analysis. Section VI discusses the results and implications for future research.

2. Policy Background

Institutions have commonly followed one of three types of grading reforms: (i) implementing grade targets; (ii) providing information on the distribution of grades in different courses to students, graduate schools, and employers, and (iii) changing the grading scale. This paper is focused on the effect of a change in the grading scale.

Grades are assigned by either a whole-letter system (e.g., A, B, C, D, F) or a plus/minus system (e.g., A, A-, B+, B, B- etc.). A change in grading scale from whole-grade grading scale to plus/minus grading has been adopted in US universities as a way to reduce grade inflation, enhance better differentiation among students, and increase student motivation. [4] Between 1992 and 2002, the percentage of institutions that reported using a plus/minus grading scheme grew by

[4] On trends towards plus/minus grading, see Bressette (2002); Quann (1987) and Riley et al., (1996).

20 percentage points, which indicates that institutions have been increasingly adopting this grading system. [5]

Two universities in Florida, the University of Central Florida (UCF) and University of South Florida (USF), changed their grading scale from whole-letter to plus/minus grades.[6] At the beginning of Fall 2000, USF implemented a plus/minus grading system that aimed to more clearly reflect the academic achievement of individual students in their courses. Similarly, but a year later, in the Fall 2001, UCF implemented the same grading scale. According to the UCF Undergraduate Policy and Curriculum Committee, their previous whole-letter grade grading system (A, B, C, D, F) placed students with widely different achievements with the same grade. This is why they decided to adopt a more continuous grading scale. While this option was not mandatory for both universities, faculty, particularly at UCF, expressed a preference for the greater exactness plus and minus grades provide.[7]

Drawing on prior literature, I hypothesize that a more refined grading scale would reduce grade differentials between STEM and non-STEM courses by potentially reducing grade inflation in non-STEM fields and/or increasing student effort in STEM fields. A reason that has motivated a more refined grading scale is that it tends to decelerate grade inflation by discouraging professors from bumping grades up to the next grade. With the possibility of pluses and minuses, rather than giving an A instead of a B, a professor may give the student a B+ or an A-. In addition to this, in Florida the distribution of numeric grades within the A scale is not symmetric. The A+ point value has the same 4.0 points as does the grade of A. Because non-

[5] In the 2005 American Association of College Registrars and Admissions Officers (AACRAO) report on grading trends, 56 percent of responding institutions indicated the use of plus/minus a part of their regularly reported grades.
[6] The other eight universities already had a plus/minus grading scale, except from the Florida Agricultural and Mechanical University that used a whole-letter grading scale during the period of analysis.
[7] It is worth nothing, as depicted in the table below, that the grade point value assigned to plus/minus grades was slightly different between these two universities (See Appendix Table A3).

STEM fields have a higher proportion of A's than STEM fields, the proportion of A's downgraded to A- is expected to be relatively higher in non-STEM courses and this would depress non-STEM GPAs.[8]

Plus/minus grades may also provide students with greater motivation to do better, particularly if the student intends to major in STEM upon initial enrollment. With the change in the grading scale, students interested in STEM, who on average receive lower grades than in non-STEM courses, may have a greater chance to improve their grades by allocating more effort to STEM courses. Also, students who are at the margin of earning a higher grade may feel more motivated to achieve it. Under a whole-letter system a student running a B in a STEM course may feel that an A is almost impossible to reach, while a B+ or even an A- is within her grasp with an additional effort. Similarly, a student satisfied with a B in a STEM course, may be slacking off since the risk of falling to a C is relatively low with whole-letter grades, whereas with plus/minus grades a B- might be a real possibility.[9]

Finally, though not the intent of the policy, the change in the grading scale could lead to an increased grade inflation if professors' goal is to give as high grades as possible while still maintaining distinctions among students. Yet, professors of STEM and non-STEM courses might respond differently to this policy, depending on their ability to better differentiate students or to justify higher grades.

[8] If plus and minus grades are included for each letter grade available, then theoretically, there is no reason to believe that there should be any change in average GPAs.
[9] Previous research has studied the effect of grades and grading standards on student effort. While Grant and Green (2013) found that students do not increase effort to raise exam scores even if it could be the difference between failing or passing the course, Grove and Wasserman (2006) found that assigning grade incentives to homework assignments increases freshmen's mean grades by about half a letter grade. There is also evidence that suggest that tougher grading standard increases student effort (Betts & Grogger, 2003; Figlio & Lucas, 2004) by motivating students toward obtaining higher letter grades (Main & Ost, 2014).

3. Data and Descriptive Statistics

This paper uses data from the Florida Department of Education's K-20 Education Data Warehouse (K-20 EDW), an integrated longitudinal dataset that covers all public school students in the state of Florida. Florida's student data-tracking system is very comprehensive and allows me to control for demographic characteristics (including if the student qualified for free lunch), SAT/ACT scores, [10] and degree and major intentions at first enrollment. This administrative data also include college characteristics such as term-by-term college enrollment (credits attempted/completed, term and cumulative GPA, and major), transcript and degree information for all post-secondary students at public institutions in Florida. The data were supplied to the author by the Florida Department of Education. [11]

The benefit of using this information is that it includes detailed information pertaining to the college experience and pre-collegiate human capital. Thus, I can observe students in each semester and year and evaluate how a change in the grading scale might differentially affect average STEM and non-STEM grades during college. These data also allow me to identify students pursuing STEM fields who take STEM and non-STEM courses in college, so as to explore how college grades might differentially affect STEM graduation/major choice. STEM majors are classified using the 2011 NCES list which in turn used a U.S. immigration and Customs Enforcement (ICE) list of designated STEM degree programs.[12] Finally, STEM courses

[10] Unfortunately, 44 percent of my sample from the 1996/07 cohort has missing values for high school characteristics such as grades and units in high school courses.
[11] This dataset does not include students pursuing a BA degree in a private institution or outside the State of Florida. I did not count on the National Student Clearinghouse that tracks college attendance outside the state of Florida as well as any private college enrollment in Florida.
[12] The ICE's list includes the instructional programs of interest to the analysis (mathematics, natural sciences, engineering/engineering technologies; computer/information sciences). STEM instructional programs were then classified into six STEM fields: computer and information sciences; engineering and engineering technologies; biological and biometrical sciences; mathematics and statistics; physical sciences; and science technologies (ICE, 2010). Non-STEM majors include all fields that are not STEM fields as well as general studies, undeclared or unknown majors.

were identified using Jacobson and Mokher (2009) classification of courses by field of concentration. See Appendix Table A1 for the list of courses used to differentiate STEM from non-STEM courses.

My sample is restricted to high school graduates who entered a 4-year public institution straight from high school for the first time in 1996, 2000 and 2003. These students were tracked forward to 2012 although the last term enrolled for most of these students is Fall 2000, Fall 2005 and Fall 2008, respectively. For these students, I have complete college transcript records. I further limit the sample to students who in May of their high school completion year were less than 21 years old, which is the baseline sample used for the descriptive analysis. Across these three cohorts I have about 59,927 observations, which is the baseline sample used for the descriptive analysis and the difference in difference analysis.

I separate colleges into institutions that changed their grading scale from whole-letter to plus/minus grades (hereafter called treatment institutions) and other 4-year public institutions (or untreated institutions). For each student and type of college, my outcome variable of interest is STEM and non-STEM graduation and major choice as well as the probability of never earned a BA.[13] STEM and non-STEM major choice was identified at two points in time: (i) at the beginning of their second year of enrollment; and (ii) at the beginning of the third year of enrollment when most undergraduates formally declare their major. STEM and non-STEM graduation outcomes are based on graduation in 6 years or less.

[13] These three outcome variables are mutually exclusive. For instance, STEM graduation is equal to one if the student graduated in STEM and equal to zero if she graduated in non-STEM, as well as if never graduated. Similarly, non-STEM graduation is equal to one if the student graduated in non-STEM and equal to zero if she graduated in STEM or never graduated. Finally, the probability of never earned a BA is one if the student never completed any BA (and 0 if graduated in any field).

Table 1a and 1b display descriptive statistics separately for the periods before and after the change in the grading scale and for those students attending treated and untreated institutions. Thirty-one percent of the sample attended treated institutions.[14] Students attending treated institutions are less likely to be female and Black, and they have lower SAT reading and math scores. Students attending treated institutions are more likely to intend to major in STEM upon first enrollment. Most differences between the two groups are stable over time. As the last column shows, [15] during the period of analysis there is a statistically significant negative effect on the proportion of students who are Hispanics and qualified for free lunch in grade 12 (also see Appendix Table A2). Students attending treated institutions are differentially trending towards better economic background. The differences between the two groups in the proportion of Hispanics and low-income students suggests that students at untreated institutions look worse on a number of dimensions that might predict grade differentials and STEM graduation/major choice. I address this potential concern in the next section.

Table 1b also shows probabilities of majoring and graduating in STEM and non-STEM, number of credits attempted in STEM and non-STEM courses, and STEM and non-STEM grades and grade differentials. For the cohort of high school graduates who entered a 4-year public institution straight from high school for the first time in 1996, those who attended treated institutions are less likely to major in STEM and graduate in 6 years of less with a STEM major than students in untreated institutions: 7 percent graduated with a STEM major by 2002, while 10 percent of students who attended an untreated institution had done so. The differences in

[14] There are 13,695 students in the pre-policy period (3,852 and 9,843 in the treatment and control groups, respectively) and 46,232 students after the grading reform, 14,724 and 31,508 students in the treatment and control groups.

[15] With only 10 institutions, I account for the small number of clusters by calculating the statistical significance relative to small sample t-distribution with 9 degrees of freedom with clustering standard errors at the institution level.

STEM major choice between these two groups increase from the second year of enrollment to the sixth year by 2.3 percentage points. For the younger cohorts of students, 2000 and 2003, the pattern is reversed: the difference in the probability of graduating with a STEM major between students attending treated and untreated institutions is no longer statistically significant. In fact, the proportion of students choosing a STEM major at the beginning of their second and third year of enrollment is now higher than that of the control group.

This table also illustrates the means in grades and first year grade differentials as well as the number of credits attempted in STEM and non-STEM courses. These college academic characteristics are focused on lower-division courses, which are mainly suitable for freshmen and sophomores, given that these courses are those that students are expected to complete in the first two years of study in their major choice, and sometimes they serve as prerequisites for upper-division courses. Grade differentials between STEM and non-STEM courses are measured using two variables: the probability of doing better in STEM (STEM GPA higher than non-STEM GPA) and the differences between STEM GPA and non-STEM GPA. For the empirical analysis, I will focus on the grades earned during the first year of enrollment.

Table 1a. Descriptive Statistics

	Before: Cohort 1996		After: Cohorts 2000 and 2003		Difference-in-Difference	
	Control	Treatment	Control	Treatment		
Demographics						
Female	0.602	0.566	0.601	0.583	0.017	
White non-Hispanic	0.624	0.698	0.603	0.693	0.017	
Black non-Hispanic	0.180	0.122	0.174	0.119	0.003	
Hispanic	0.126	0.117	0.156	0.112	-0.033	**
Asian or pacific islander	0.051	0.050	0.050	0.055	0.006	*
Other or unknown race	0.019	0.014	0.017	0.021	0.008	***
Student is a US citizen	0.936	0.956	0.943	0.960	-0.003	
Age in May of HS completion year	18.431	18.427	18.201	18.194	-0.003	
Qualified for Free Lunch in Grade 12	0.069	0.076	0.087	0.067	-0.027	**
High School Performance						
Highest SAT reading Score	552.173	536.658	552.733	541.806	4.588	
Highest SAT math Score	555.870	539.403	559.584	550.358	7.241	

Table 1b. Descriptive Statistics

	Before: Cohort 1996		After: Cohorts 2000 and 2003		Difference-in-Difference	
	Control	Treatment	Control	Treatment		
College Experience						
STEM semester 1	0.208	0.247	0.179	0.217	-0.002	
First Semester STEM GPA	2.249	1.882	2.100	2.084	0.352	**
Second Semester STEM GPA	2.479	2.031	2.419	2.340	0.370	**
First Year STEM GPA	2.273	1.881	2.189	2.153	0.355	***
First Semester Non-STEM GPA	2.500	2.404	2.438	2.546	0.203	*
Second Semester Non-STEM GPA	2.838	2.691	2.947	2.845	0.046	
First Year Non-STEM GPA	2.646	2.437	2.660	2.619	0.168	**
First Year STEM GPA>Non-STEM GPA	0.309	0.238	0.286	0.297	0.082	***
First Year GAP Diff.	-0.430	-0.583	-0.498	-0.500	0.150	***
Num STEM lower-division courses enrolled	18.591	12.016	20.526	18.038	4.087	***
Six-Year STEM credits attempted	11.492	7.960	16.545	14.231	1.217	
Six-Year STEM credits earned	10.888	7.346	15.687	13.242	1.096	
Num Non-STEM lower-division courses enrolled	26.883	19.939	27.582	25.525	4.886	**
Six-Year Non-STEM credits attempted	51.510	40.892	62.627	58.164	6.155	
Six-Year Non-STEM credits earned	56.836	51.4527	63.160	59.752	1.975	

	Before: Cohort 1996		After: Cohorts 2000 and 2003		Difference-in-Difference	
	Control	Treatment	Control	Treatment		
Graduated in 6yrs with a STEM major	0.104	0.073	0.093	0.090	0.028	
STEM year 2	0.145	0.121	0.128	0.138	0.034	*
STEM year 1	0.178	0.170	0.155	0.177	0.030	
Graduated in 6yrs a Non-STEM major	0.551	0.445	0.540	0.456	0.021	
Non-STEM year 2	0.604	0.527	0.619	0.577	0.036	**
Non-STEM year 1	0.579	0.564	0.598	0.603	0.019	
Number of Observations	9,843	3,852	31,508	14,724	59,927	

Source: Author using student-transcript-level data from Florida Department of Education
Note: Standard errors are in parentheses. Standard errors in the difference-in-difference column adjusted for clustering at the institution level. Statistical significance in the difference-in-difference column was calculated relative to the small sample t-distribution with 9 degrees of freedom.

4. Empirical Specification

In this section, I estimate the effect of changing the grading scale from whole letter grades to plus-minus grades using one cohort of students before (1996-97) and two cohorts after (2000-01 and 2002-03) this grading policy was implemented. I then compare their grade differentials between in STEM and non-STEM lower-division courses and STEM major choice and graduation outcomes over the six years following initial enrollment.

A difference-in-difference approach enables me to compare similar students over time at institutions with or without a change in grading scale. First, I estimate the effect of adopting a new grading scale on average first year grade differentials. Then, I will use the same identification strategy to estimate the reduced-form difference-in-difference equation:

$$(1)\Pr(STEM\ major_{ij=1}) = \emptyset(\alpha + \beta Instreat_{ij} * After_j + \delta InstFE_{ij} + \sigma After_i + \gamma X_{ij} + \mu_{ij}$$

where $\Pr(STEM\ major_{ij=1})$if the student i in school j chose a STEM at the beginning of her second and third year of enrollment and graduates with a STEM major in 6 years or less. $After_j$ is a dummy variable that takes the value of 0 if the student entered a 4-year public institution straight from high school for the first time in the Fall of 1996, and 1 if the student entered in the Fall of 2000 and 2003. $Instreat_{ij}$ takes the value 1 if the institution change its grading scale and

0 if the institution did not change its grading scale. $InstFE_{ij}$ is a complete set of institution fixed effects, and X_{ij} is a vector of covariates that controls for demographic and pre-college characteristics. Covariates include dummy variables for female, US citizen, those who qualified for free lunch in grade 12, age and age squared at first entry, and STEM initial major intention, SAT reading and math scores, and SAT reading and math scores squared. The coefficient of interest in this model is β, which is the difference-in-difference estimate conditional on the observable characteristics as well as on those characteristics interacted with $Instreat$ and $After$ indicators. The same equation will be estimated for non-STEM graduation/major choice outcomes as well as for the probability of never completing a BA.[16]

The critical identifying assumption in the differences-in-difference approach is that the coefficient on the interaction term from Equation (1) would be zero in the absence of the grading reform. In other words, there are fixed, time-invariant differences across students attending treated and untreated institutions and that the change in the grading policy is the only factor altering these differences over time. Pre- and post-policy cohorts of students may be different, and within each cohort, students attending treated and untreated institutions may be different; but it cannot be the case that there is something different about being a post-policy student attending a treated institution, other than the new grading scale. Trends in pre-grading policy cohorts should provide good predictors of what would have happened in the absence of this grading reform. Basically, trends in grades and STEM graduation and major outcomes do not differ pre-policy change.

To test this assumption, I estimate a version of Equation (1) with no covariates, with background characteristics as the dependent variable as shown in the last column of Table 1a. As

[16] Standard errors are robust and clustered at institution level. Results with unclustered standard errors are less conservative.

previously noted, it seems that students attending treated institutions are relatively getting slightly better in SAT scores and economic background. If the interaction terms have a significant effect on these observable characteristics, it raises the concern that there may be unobservable differences too, which could cofound my impact estimates. I find significant differences in the proportion of students who are Hispanics and qualified for free lunch in grade 12. All these background characteristics are included as controls in the analysis, but it is worth noting that these variables only explain 17 percent of the post-policy change in the first year GPA in STEM lower-division courses (See Appendix Table A4). It may also be unfeasible for these differences to differentially affect the GPA differentials between STEM and non-STEM courses.

Students' assignment would invalidate the parallel trends assumption of the difference-in-difference approach; however, it is very unlikely to think that those who are more interested in graduating in STEM may have chosen an institution that changed its grading scale. Even if students' college choice were potentially endogenous, the difference-in-differences framework would address this problem to the extent that differences between the students attending treated and untreated institutions show up in the baseline levels of STEM major intention. As shown in Table 1b, these differences are insignificant and negative, which suggests that control institutions are relatively attracting more students who intend to major in STEM.

Yet, differences in changes in STEM graduation/major choice outcomes due to unobserved differences between the students of both treated and untreated institutions might invalidate the parallel trend assumption. Even though I only have data for three cohorts, the pre-policy cohort and two cohorts after, I used the overall number of STEM degrees awarded in the State University System of Florida (see Figure A1) and Appendix Figure A8-Figure A10 to assess if there were differential pre-treatment trends in STEM graduation rates. In addition to this, I provide details about trends in grades during the first two semesters of enrollment for the three cohorts (see Appendix Figure A2-Figure A7).[17] Overall, the trends in STEM and non-STEM grades (and STEM and non-STEM grade distributions) as well as in STEM graduation strongly changed after the new grading scale was implemented among treated institutions to catch up with that of untreated institutions.

Finally, the parallel trend assumption also implies that in the absence of the policy change students attending treated institutions would have been exposed to the same institutional policies or environment as the students attending untreated institutions. Nevertheless, it is unlikely that treated institutions and untreated institutions systematically implemented policies oriented to differentially affect graduation outcomes between STEM and non-STEM departments at the same time.

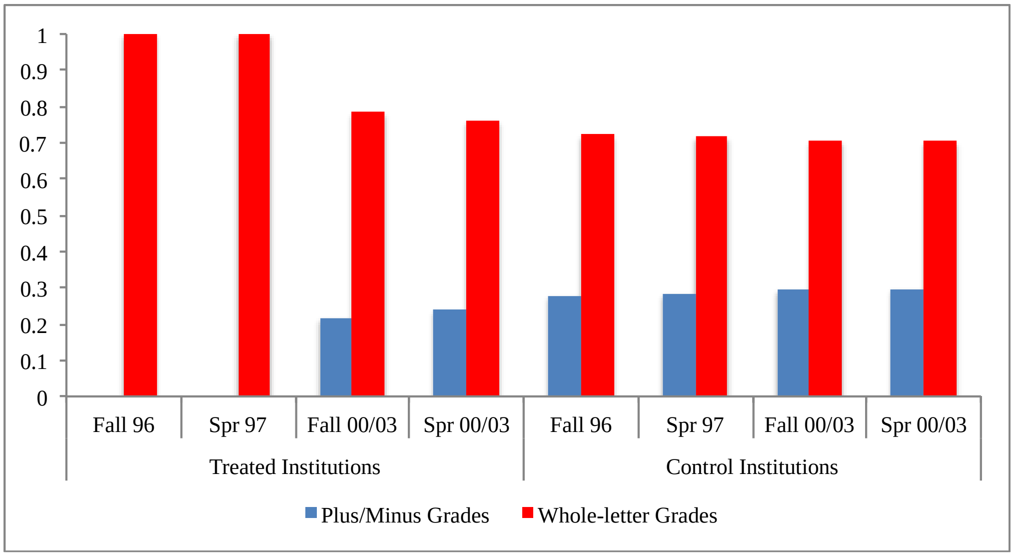
5. Empirical Results

Figure 1 shows the percentage of students who obtained whole and plus/minus letters in lower-and upper-division courses before and after the grade policy among the two institutions that implemented a plus/minus scale and the remaining eight institutions. On average, before the change in the grading scale, students who attend treated institutions during their first year of enrollment do not receive plus-minus grades, while almost 30 percent of students in untreated

[17] Pre-policy change treated and untreated institutions were using a different grading scale, which means that post-policy change I should expect trends to be more comparable. Before the policy, the gap in the percentage of A and C/F grades in STEM lower-division courses between students attending treated and untreated institutions is around 10 to 20 percentage points; after the policy, this gap was gradually closed. For non-STEM courses, the gap in the percentage of A and C/F grades was insignificant before and after the policy. Students attending untreated institutions should not be affected by the change in the grading scale, and indeed there is no shift in grades and the fraction of STEM and non-STEM graduates among this group after the policy. Prior to the grading reform, there is a downward (upward) trend in the percentage of A (C/F) grades in STEM courses among students attending treated institutions. But after the reform, this trend reverses.

institutions obtained plus/minus grades. After the adoption of the plus/minus grading system, treated institutions significantly increased the percentage of plus/minus grades assigned to GPAs by approximately 23 percentage points.

Figure 1 Pre-and-post Policy Percentage of Whole and Plus/Minus Grades by Enrollment Term

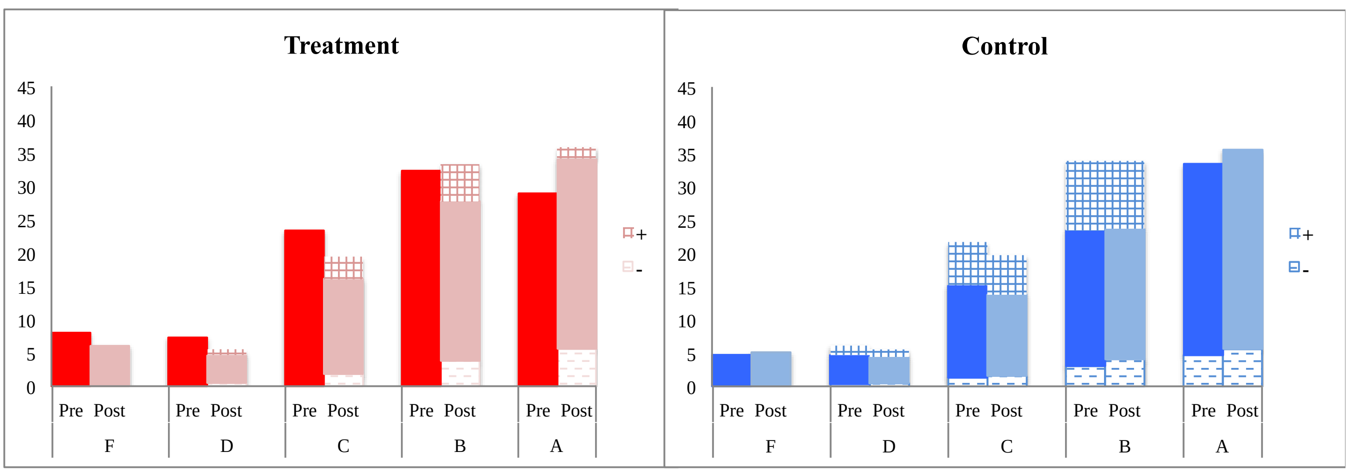


Source: Author using student-transcript-level data from Florida Department of Education
Note: The figure displays percentages of whole and plus/minus grades in lower-and upper-division courses during the Fall 1996 and Spring 1997, before the change in the grading scale, and the Fall 2000/03 and Spring 2000/03, after the grading policy was in effect.

Although a change in the grading scale does not seem to affect overall average grades, it has had a differential effect on grade distribution between STEM and non-STEM departments and particularly in treated institutions. Figures 2a-2c provide actual grade distributions in the A, B, C and D/F ranges for all courses taken as well as for those taken in STEM and non-STEM departments. An initial glance at data from three cohorts of students who first enrolled in 1996 (pre-policy change), and in 2000 and 2003 (post-policy change) suggests that the proportion of students attending treated institutions who received more than B+ in all lower-division courses has increased by 24 percent post-policy change. Similarly, the proportion of those who received grades less than B- decreased by 20 percent. This pattern is also manifested in untreated institutions but with a change of 2 percentage points (about 5 percent growth) in the proportion of grades at the top and bottom of the grade distribution.

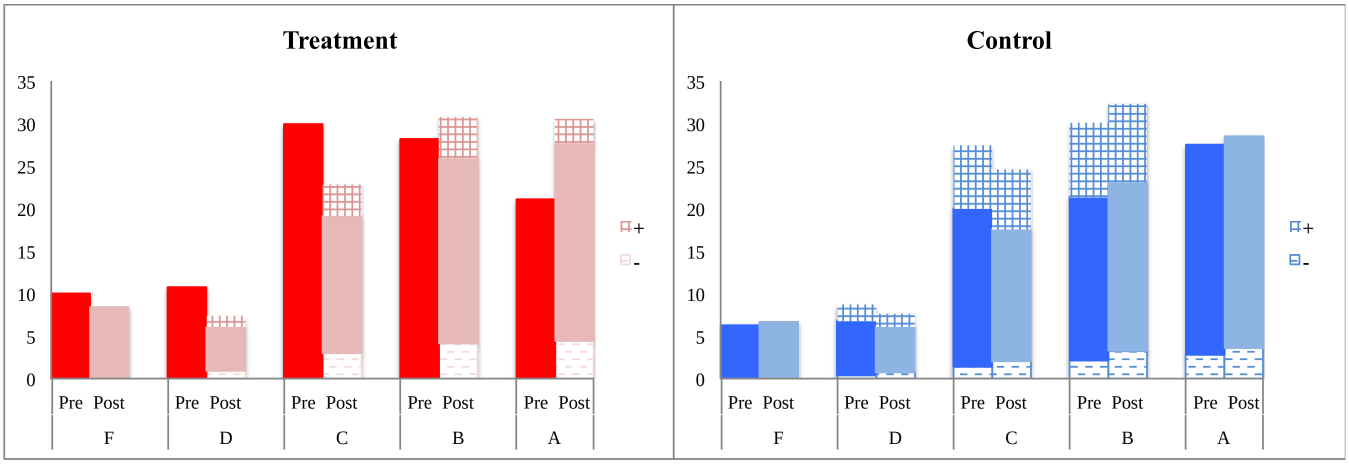
After the plus/minus grading scale was implemented in treated institutions, the grade distribution shifted higher for STEM courses than for non-STEM courses. In fact, the grade distribution for non-STEM courses followed closely the pattern of the overall distribution of grades. For these courses the change in the proportion of grades at the top and bottom of the grade distribution was about the same, 7 percentage points. In contrast, for STEM courses the proportion of students receiving grades higher than B+ has increased from 21 percent to 31 percent and the proportion receiving less than B- has decreased from 51 percent to 39 percent. Moreover, comparing the pre-and post-policy change periods, the proportion of grades earned in STEM courses at the top of the grading scale changed at a higher rate than the proportion at the bottom of the grading scale (45 percent increase versus 24 percent decrease, respectively). Therefore, there is evidence that the grade distribution strongly changed in treated institutions after the change in the grading scale and especially for students taking STEM courses.

Figure 2a. Distribution of Grades by Treatment Status



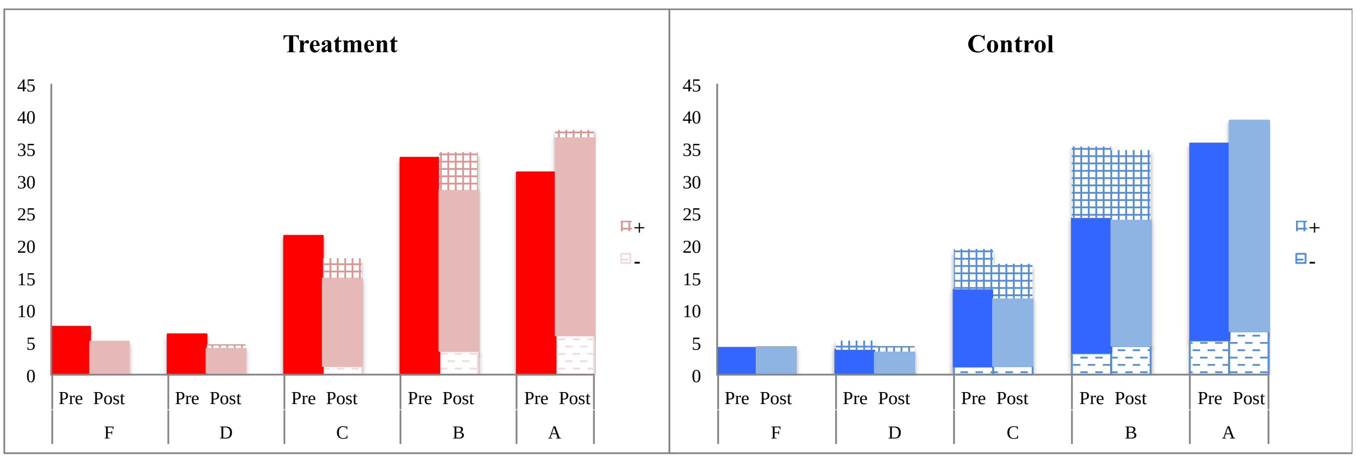
Source: Author using student-transcript-level data from Florida Department of Education
Note: These figures display the share if student-course observations that were assigned a letter grade before and after the change in the grading scale by treatment status. Grades for lower-division courses taken during the first year of enrollment. The percentage of whole and plus/minus grades are stacked in the graph.

Figure 2b. Distribution of STEM Grades by Treatment Status



Source: Author using student-transcript-level data from Florida Department of Education.
Note: These figures display the share if student-course observations that were assigned a letter grade before and after the change in the grading scale by treatment status. Grades for lower-division STEM courses taken during the first year of enrollment. The percentage of whole and plus/minus grades are stacked in the graph.

Figure 2c. Distribution of Non-STEM Grades by Treatment Status



Source: Author using student-transcript-level data from Florida Department of Education
Note: These figures display the share if student-course observations that were assigned a letter grade before and after the change in the grading scale by treatment status. Grades for lower-division non-STEM courses taken during the first year of enrollment. The percentage of whole and plus/minus grades are stacked in the graph.

For students attending untreated institutions, the change in the grade distribution for STEM courses was rather small, and the distribution of grades pre-and post-policy change is very similar. The proportion of students who received less than B- has risen by 3 percentage points post-policy change, while the proportion of students who received more than B+ only increased by 1 percentage point. Moreover, this change was 3 percentage points smaller than that in the proportion of students receiving above B+ in non-STEM courses.

As a result, after the change in the grading scale the difference in average grades between treated and untreated institutions substantially decreased as well as the grade differentials between STEM and non-STEM departments in treated institutions. Figure 3a and 3b show what happened to average STEM and non-STEM GPAs in treated and untreated institutions before and after the change in the grading scale. STEM grades actually show a substantial catch up between treated and untreated institutions after the change in the grading scale was in effect. Even though both STEM and non-STEM grades rose after the change in the grading scale, this change was much bigger among treated institutions and STEM departments and especially during the first year of enrollment. In the first semester of 1996 and the first semester of 2000 and 2003, the average STEM grades climbed from around 2.42 to 2.75 among treated institutions, an increase of more than 13 percent.[18]

Figure 4a shows that differences in grades between STEM and non-STEM lower-division courses among treated institutions was reduced during the first two semesters of enrollment post-policy change. Among untreated institutions the average differences in grades between STEM and non-STEM courses practically did not change after the change in the grading scale. Similarly, on average the proportion of STEM grades that are higher than non-STEM grades increased by 11 percentage points between the first year of enrollment pre-policy change and post-policy change (2000 and 2003), as shown in Figure 4b.

Appendix Figure 11a and 11b show how the distribution of the differences in grades between STEM and non-STEM courses changes during the first year of enrollment before and after the new grading scale and between students attending treated and untreated institutions. While the

[18] In contrast, the average non-STEM grades were around 3.15 in the first semester after the change in the grading scale, an increase of 4 percent from the first semester of 1996.

distribution of grade differentials for students attending treated institutions shifts upward towards reducing the first year GPA differentials, the distribution for those attending untreated institutions shifts backwards, which implies a secular grade inflation.

The decrease in grade differentials between STEM and non-STEM courses by a disproportional increase in STEM grades does not completely support my hypothesis. A more refined grade scale seems to accelerate grade inflation in STEM departments, leaving grades in non-STEM departments mostly unaffected. Yet, my results provide support for the possibility of equalizing grades using this grading policy, although it is not totally obvious why STEM grades increased. Professors' responses to the change in the grading scale together with students' decisions to allocate more effort to STEM versus non-STEM courses might explain how grades differentials are reduced under this policy. The latter hypothesis is hard to prove because I would have to measure the effect of the policy on students who are at the margin of achieving a higher grade and then check if the shift in the distribution of STEM grades stems from those who could have been potentially motivated to exert more effort. Without having data on numerical scores or study time, it is hard to assess this hypothesis.

Figure 3a. STEM Grades by Enrollment Term

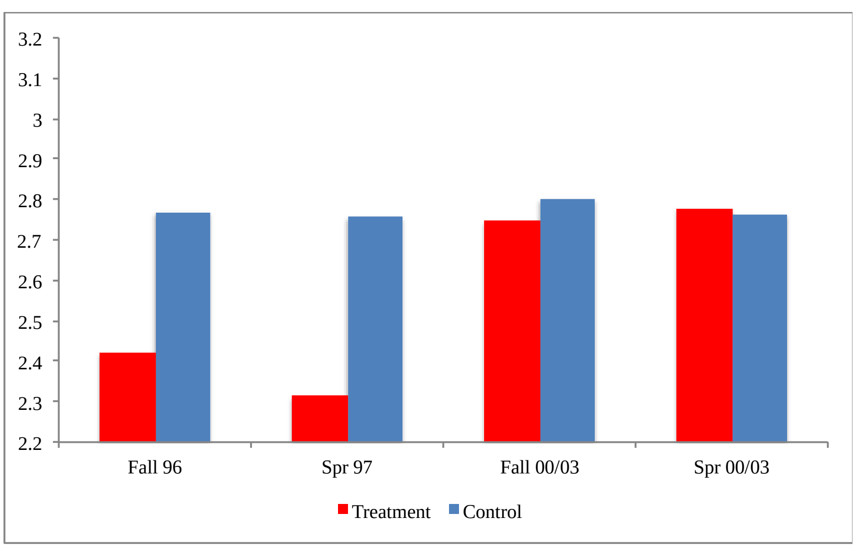
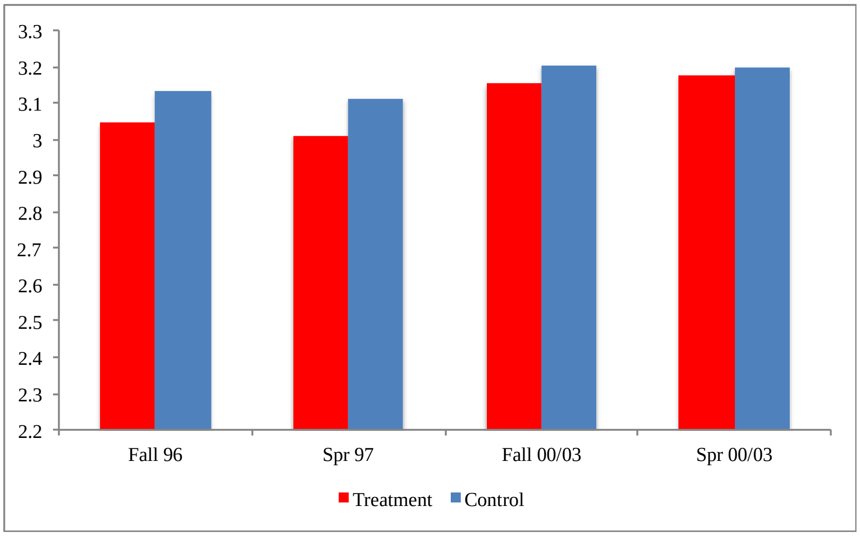


Figure 3b. Non-STEM Grades by Enrollment Term



Source: Author using student-transcript-level data from Florida Department of Education
Note: The figure displays mean Non-STEM grade in lower-division courses during the Fall 1996 and Spring 1997, before the change in the grading scale, and the Fall 2000/03 and Spring 2000/03, after the grading policy was in effect.

Figure 4a. Percentage of STEM GPA higher than non-STEM GPA by Enrollment Term

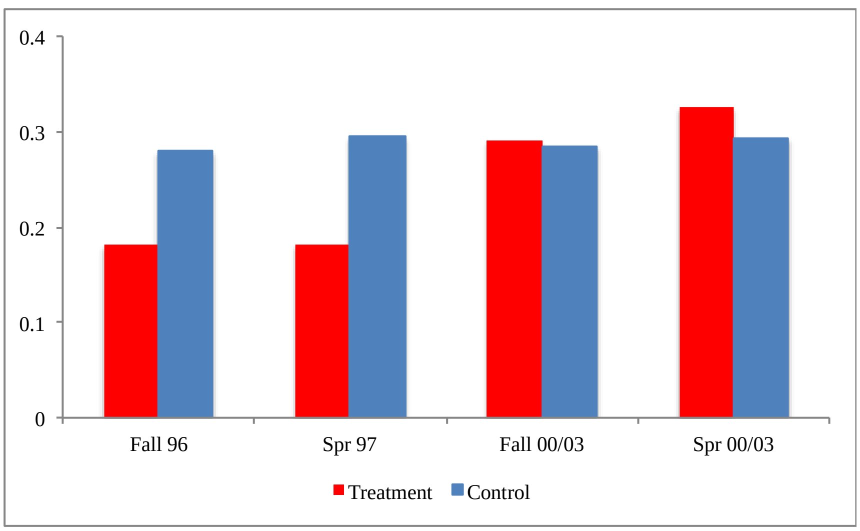
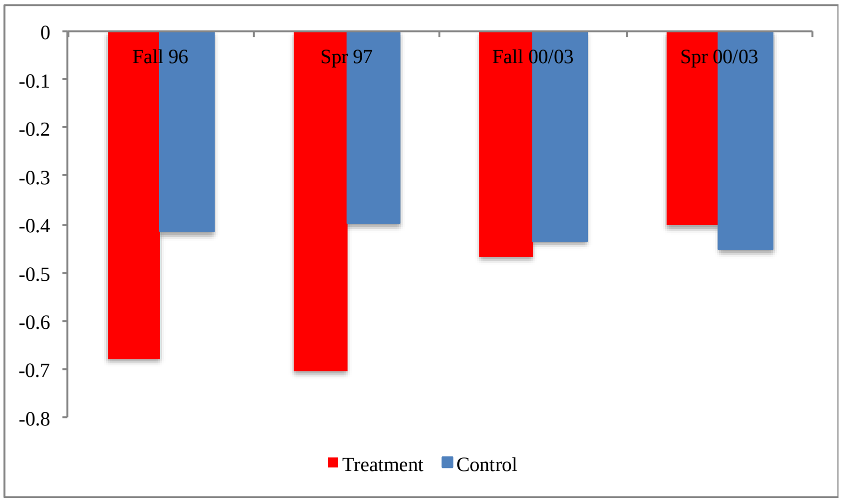


Figure 4b. Difference between STEM GPA and non-STEM GPA by Enrollment Term



Source: Author using student-transcript-level data from Florida Department of Education
Note: The figure displays average difference between STEM grade and Non-STEM grade in lower-division courses during the Fall 1996 and Spring 1997, before the change in the grading scale, and the Fall 2000/03 and Spring 2000/03, after the grading policy was in effect.

My findings could also be potentially explained by professors' preferences for "inflating" STEM grades to better differentiate students. In both STEM and non-STEM fields, professors are under pressure to give higher grades. However, professors in STEM fields might care more about (or find it easier to justify) having a grade distribution at all to better differentiate students, while professors in non-STEM fields might be more comfortable with the idea of compressing grades at the top. Moreover, the cost of grading associated with student complaints while keeping fair distinctions among students may be easier in STEM than in non-STEM courses. This could be particularly the case if relatively more students in STEM are at the margin of achieving higher grades than in non-STEM fields. Consequently, plus/minus grades allow STEM professors to inflate grades like in other fields, but without losing distinctions between students.[19]

[19] In non-STEM courses essays are more subjective than formulas and problem sets in STEM courses, and hence tougher to evaluate and relatively more subjected to students' complaints. Moreover, non-STEM professors might let students rewrite their papers and then give them new grades, which is not a common practice with STEM fields.

Even though future research is needed to investigate why professors in STEM departments are more sensitive to this grading policy, it seems that professors in STEM are making relatively more distinctions among students by using a higher percentage of plus/minus grades than in non-STEM courses. After the change in the grading policy, plus/minus markers were 5 percentage points more used in STEM than in non-STEM departments.[20]

The first and second columns of Table 2 show "first stage" results from estimating Equation (1) and find that after controlling for the after years, whether or not the student was enrolled in an institution that change the grading scale, institutional fixed effects, and student background characteristics, there is a significant relationship between the interaction term and the two measures of grade differentials (i.e. STEM GPA higher than non-STEM GPA and the difference in STEM GPA and non-STEM GPA). Model 1 reports the difference-in-difference estimate with institutional fixed effects. Model 2 controls for math and reading SAT scores (and SAT scores squared). It also includes the same dummy variables included in Equation (1), and it is reassuring that the overall pattern of results remains consistent. Having attended an institution that changed its grading scale leads to 8 more percentage points in the probability of earning a higher cumulative STEM GPA than non-STEM GPA after the new grading scale was in effect. On average, before the new grading scale was enacted, about 24 percent of students who attended treated institutions did better in STEM than in non-STEM during their first year of enrollment, so this percentage grew about 30 percent.

Similarly, the impact of the change in the grading scale on the difference between STEM and non-STEM GPAs during the first year of enrollment is about 0.152, a relative increase of about a

[20] This estimate only considers courses that actually change the grading scale given that not all faculty members in treated institutions used plus/minus markers after the policy. Without restricting my sample, this estimate is 2 percentage points less.

sixth of a grade differential in treated institutions. It is worth noting that the effect on the difference in grades is embedding a triple difference-in-difference, which assumes that there is no shock after the change in the grading scale that had differentially affected grade differentials of students who attended treated institutions. This effect is mainly explained by the impact of the policy on STEM grades, which is about 0.28 grade points.

The third and fourth columns of Table 2 present the main reduced-form results. For each outcome, Model 1 shows estimates from the most basic difference-in-difference model, with the interaction term, and fixed effects for institution, but no covariates. Model 2 adds covariates and improves the precision of my impact estimates by significantly reducing standard errors without altering the overall pattern of results. The results from Model 2 indicate large, positive and statistically significant impacts on STEM graduation/major choice outcomes and very small and insignificant results on non-STEM graduation/major choice as well as on the probability on never completing a BA. However, these results suggest that the relative increase in the probability of majoring in STEM is mainly explained by a reduction in the probability of dropping out of college. In fact, results also indicate large and statistically significant impacts on the number of lower-division courses enrolled in 6 years or less in both STEM and non-STEM fields due to a relatively lower dropout rate.

Focusing on Model 2 from Table 2 column (4), after the grading policy change, students were about 3 percentage points more likely to graduate in STEM in 6 years or less or to choose a STEM major at the beginning of their second and third year of enrollment in the treated institutions. On average before the implementation of the new grading scale, about 7 percent of students attending treated institutions graduated in STEM, hence my impact estimate represents an increase of about 43 percent. Results are mostly significant among STEM graduation/major choice outcomes. The results in Table 2 for non-STEM outcomes are noisier and indicate no statistically significant impacts of the change in the grading scale on non-STEM and BA completion outcomes.

A similar recent study that also relies on a difference-in-differences methodology supports these findings. Butcher et al. (2014) evaluate the effect of an anti-grade-inflation policy that reduces grades in high-grading departments and find that students' choices about courses and majors are sensitive to grades. Their results suggest that majors declined in high-grading departments by about 30 percent while the fraction of a graduating class majoring in economics (low-grading department) increased. Although the change in the grading scale positively affected grades in low-grading departments (STEM fields), the mechanisms through which this policy works are similar to those studied in Butcher et al. (2014). My results can be scaled for comparison with the -0.17 grade point impact found in their treated departments. Thus, the 0.28 grade point impact found on STEM grades would correspond with a 50 percent increase in the probability of graduating in STEM. This estimate is not significantly bigger than what I found.

Gender

Table 3 and 4 show results using the preferred specification (Model 2), estimated separately for women and men. The first column of Table 3 and Table 4 suggests that a change in the grading scale has a positive and significant effect on the reduction of first year GPA differentials between STEM and non-STEM courses for women and men at similar rates. This is consistent with the effects on STEM graduation/major choice outcomes, which are not significantly different between women and men. Proportionally, after the change in the grading scale both women and men are relatively increasing their graduation rates in STEM by about 50 percent.

Table 2. First Stage and Reduced Form Difference-in-Difference Estimates of the Effect of the Change in the Grading Scale between STEM and non-STEM courses

	(1)	(2)	(3)	(4)
	First Stage		**Reduced Form**	
VARIABLES	**Model 1**	**Model 2**	**Model 1**	**Model 2**
First Year STEM GPA> Non-STEM GPA	0.086	0.080		
	(0.012)***	(0.012)***		
First Year GPA Diff.	0.162	0.152		
	(0.039)***	(0.036)***		
First Year STEM GPA	0.299	0.281		
	(0.098)**	(0.086)***		
First Year Non-STEM GPA	0.122	0.105		
	(0.078)	(0.072)		
STEM Major year 2			0.020	0.030
			(0.021)	(0.005)***
STEM Major year 3			0.026	0.033
			(0.015)	(0.006)***
Graduating in 6yrs with a STEM major			0.020	0.025
			(0.015)	(0.006)***
STEM credits attempted in six years or less			0.523	0.820
			(0.943)	(1.326)
Number of STEM courses enrolled in 6yrs or less			3.610	3.685
			(0.896)***	(0.815)***
Non-STEM Major year 2			0.011	0.001
			(0.026)	(0.015)
Non-STEM Major year 3			0.022	0.014
			(0.013)	(0.024)
Graduating in 6yrs with a Non-STEM major			-0.000	-0.008
			(0.055)	(0.064)
Non-STEM credits attempted in six years or less			4.127	3.741
			(4.008)	(4.410)
Number of Non-STEM courses enrolled in 6yrs or less			4.004	3.970
			(1.747)**	(1.666)**
Did not earn a BA in 6yrs			-0.020	-0.017
			(0.068)	(0.067)

Note: This table shows results for the first stage and reduced form estimates. Model 1 only includes institution fixed effects. Model 2 controls for covariates and institution fixed effects. Variables are imputed to zero if missing; missing data flags are included for all variables with missing data. Standard errors clustered at institution level. Sample size for the reduced form of Model 2 is 59,813; sample size for the first stage and IV identification is 48,060 due to missing values in grade differentials. The covariates used for Model 2 include dummy variables for female, black, Hispanic, Asian, other race, US citizen, those who qualified for free lunch in grade 12, age and age squared at first entry, and STEM initial major intention, SAT reading and math scores, SAT reading and math scores squared, number of credits attempted in STEM and non-STEM lower-division courses during the first year of enrollment. IV results are only shown for Model 2 using first year GPA differentials as instrumented variable. My sample is restricted to high school graduates who entered a 4-year public institution straight from high school for the first time in 1996, 2000 and 2003.

As shown in column (2) of Table 3 and Table 4, the estimated drop in grade differentials in treated institutions is larger and statistically significant for women whereas the relative increase in first year STEM GPA is larger for men. The impact of the policy on graduating in STEM is around 3 percentage points for women and 5 percentage points for men. Interestingly, for women the positive effects on STEM graduation/major choice outcomes seem to be mainly explained by a reduction in the probability of leaving STEM by switching into non-STEM fields, whereas for men these effects accounted for primarily a decrease in the probability of leaving STEM by dropping out of college. This could suggest that for women grade differentials are perceived as a signal of relative ability in STEM versus non-STEM courses, and therefore they might be discouraged from pursuing STEM because they think they can do better in non-STEM fields. In contrast, in the case of men grade differentials may be more perceived as a signal of absolute ability or of their capacity to succeed in college irrespective of how better or worse they are in non-STEM fields.

These results do not necessarily contradict my hypothesis, but suggest that the differential effect of the policy between women and men is complex. If anything, men seemed to be more sensitive to the grade signal than women, which is counter to my prior hypothesis. If women value grades and benefit from higher-grades more than men, the effect of the policy on improving STEM graduation and major choice was expected to be higher for women. However, the fact that women mainly increase their STEM graduation rates by reducing their probability of graduation in non-STEM suggest that women may value relative grades more than men. If women tend to value higher grades that are “easier” to achieve in non-STEM courses, then a larger relative reduction in grade differentials might explain why women are more discouraged from graduating in non-STEM than their peers.

Table 3. Gender Subgroup Analysis: First Stage and Reduced Form Difference-in-Difference Estimates of the Effect of the Change in the Grading Scale for Females

VARIABLES	(1) First Stage	(2) Reduced Form
First Year STEM GPA> Non-STEM GPA	0.068	
	(0.024)**	
First Year GPA Diff.	0.173	
	(0.066)**	
First Year STEM GPA	0.277	
	(0.118)**	
First Year Non-STEM GPA	0.086	
	(0.066)	
STEM Major year 2		0.025
		(0.008)**
STEM Major year 3		0.029
		(0.009)**
Graduating in 6yrs with a STEM major		0.026
		(0.006)***
STEM credits attempted in six years or less		0.226
		(1.783)
Number of STEM courses enrolled in 6yrs or less		3.232
		(0.807)***
Non-STEM Major year 2		0.005
		(0.010)
Non-STEM Major year 3		0.001
		(0.027)
Graduating in 6yrs with a Non-STEM major		-0.025
		(0.067)
Non-STEM credits attempted in six years or less		4.439
		(4.601)
Number of Non-STEM courses enrolled in 6yrs or less		3.735
		(1.445)**
Did not earn a BA in 6yrs		-0.000
		(0.064)

Source: Author using student-transcript-level data from Florida Department of Education
Note: This table shows results for the first stage and reduced form estimates. These models include controls for covariates and institution fixed effects. Variables are imputed to zero if missing; missing data flags are included for all variables with missing data. Standard errors clustered at institution level. Sample size for the reduced form is 35,573. The covariates used these models include in dummy variables for black, Hispanic, Asian, other race, US citizen, those who qualified for free lunch in grade 12, age and age squared at first entry, and STEM initial major intention, SAT reading and math scores, SAT reading and math scores squared. My sample is restricted to high school graduates who entered a 4-year public institution straight from high school for the first time in 1996, 2000 and 2003.

Table 4. Gender Subgroup Analysis: First Stage and Reduced Form Difference-in-Difference Estimates of the Effect of the Change in the Grading Scale for Males

VARIABLES	(1) First Stage	(2) Reduced Form
First Year STEM GPA> Non-STEM GPA	0.068	
	(0.024)**	
First Year GPA Diff.	0.173	
	(0.066)**	
First Year STEM GPA	0.296	
	(0.049)***	
First Year Non-STEM GPA	0.132	
	(0.088)	
STEM Major year 2		0.049
		(0.016)**
STEM Major year 3		0.061
		(0.013)***
Graduating in 6yrs with a STEM major		0.049
		(0.015)**
STEM credits attempted in six years or less		1.402
		(1.152)
Number of STEM courses enrolled in 6yrs or less		4.288
		(0.917)***
Non-STEM Major year 2		-0.019
		(0.021)
Non-STEM Major year 3		0.010
		(0.020)
Graduating in 6yrs with a Non-STEM major		-0.012
		(0.059)
Non-STEM credits attempted in six years or less		2.895
		(4.367)
Number of Non-STEM courses enrolled in 6yrs or less		4.273
		(1.985)*
Did not earn a BA in 6yrs		-0.037
		(0.073)

Source: Author using student-transcript-level data from Florida Department of Education
Note: This table shows results for the first stage and reduced form estimates. These models include controls for covariates and institution fixed effects. Variables are imputed to zero if missing; missing data flags are included for all variables with missing data. Standard errors clustered at institution level. Sample size for the reduced form is 24,240; sample size for the first stage. The covariates used these models include in dummy variables for black, Hispanic, Asian, other race, US citizen, those who qualified for free lunch in grade 12, age and age squared at first entry, and STEM initial major intention, SAT reading and math scores, SAT reading and math scores squared. My sample is restricted to high school graduates who entered a 4-year public institution straight from high school for the first time in 1996, 2000 and 2003.

To explore this further, I evaluate what part of the grade distribution was mostly affected by the change in the grading scale so as to explore why men were more responsive to this policy. If relatively more men are at the margin of dropping out of college (or at the bottom of the grading distribution), then these overall impacts mask important patterns of heterogeneity. Figure A12 and Figure A13 present the distribution of grades by gender for both treatment and control groups before and after the change in the grading scale. Prior to the policy change men attending treated institutions earned a significantly higher proportion of D/F grades in STEM courses than women. Table 5 shows results for the effect of the policy on the distribution of grades for women and men. There is a higher reduction in the percentage of D/F grades for men than for women, 6 percentage points versus 3 percentage points, and these differences are statistically significant. This suggests that men were more influenced to persist in STEM by not failing STEM courses, which might explain why there is a higher reduction in the probability of dropping out of college for men than for women. The fact that women value higher-grades more than men and men were relatively more affected at the bottom of their grade distribution might explain why overall impact estimates are higher for men.

Minority Status

Table 6 and Table 7 show that for minorities and non-minorities a change in the grading scale has a similar positive and significant effect on the reduction of first year GPA differentials between STEM and non-STEM courses. Impact estimates for STEM graduation/major choice are generally larger for minorities, with a higher probability of choosing and graduating in STEM. The overall pattern of results, however, is noisier for minorities, and the differences in means between both groups are insignificant. However, proportionally, after the change in the grading scale racial minorities had a larger increase in their STEM graduation rates than non-racial minorities, 67 percent versus 44 percent.

Table 5. First Stage Difference-in-Difference Estimates of the Effect of the Change in the Grading Scale on Grade Distribution

	(1)	(2)	(3)	(4)	(5)
VARIABLES	**ALL**	**Female**	**Male**	**Minority**	**Non-Minority**
ALL GRADES					
A	0.022	0.014	0.036	0.012	0.024
	(0.026)	(0.025)	(0.026)	(0.022)	(0.029)
B	0.012	0.011	0.010	0.008	0.013
	(0.009)	(0.008)	(0.011)	(0.008)	(0.012)
C	-0.010	-0.009	-0.013	-0.016	-0.007
	(0.014)	(0.014)	(0.014)	(0.019)	(0.014)
D/F	-0.024	-0.016	-0.033	-0.003	-0.030
	(0.012)*	(0.010)	(0.015)*	(0.020)	(0.010)**
STEM Grades					
A	0.095	0.086	0.111	0.072	0.103
	(0.019)***	(0.018)***	(0.018)***	(0.015)***	(0.024)***
B	0.020	0.026	0.011	0.076	0.000
	(0.019)	(0.018)	(0.029)	(0.028)**	(0.015)
C	-0.069	-0.082	-0.053	-0.140	-0.044
	(0.029)**	(0.022)***	(0.047)	(0.029)**	(0.025)
D/F	-0.046	-0.029	-0.069	-0.007	-0.059
	(0.022)*	(0.019)	(0.029)**	(0.027)	(0.021)**
Non-STEM Grades					
A	0.005	0.008	0.006	0.002	0.004
	(0.031)	(0.031)	(0.031)	(0.029)	(0.033)
B	0.018	0.011	0.024	-0.006	0.027
	(0.016)	(0.016)	(0.016)	(0.018)	(0.020)
C	-0.003	-0.004	-0.003	0.009	-0.006
	(0.012)	(0.013)	(0.011)	(0.020)	(0.010)
D/F	-0.020	-0.015	-0.026	-0.005	-0.024
	(0.011)	(0.009)	(0.015)	(0.019)	(0.009)**

Source: Author using student-transcript-level data from Florida Department of Education
Note: This table shows results for first stage estimates. Estimates are shown for Model 2 which controls for covariates and institution fixed effects. Variables are imputed to zero if missing; missing data flags are included for all variables with missing data. Standard errors clustered at institution level. The covariates used for the "ALL" specification include in dummy variables for female, black, Hispanic, Asian, other race, US citizen, those who qualified for free lunch in grade 12, age and age squared at first entry, and STEM initial major intention, SAT reading and math scores, SAT reading and math scores squared. My sample is restricted to high school graduates who entered a 4-year public institution straight from high school for the first time in 1996, 2000 and 2003.

Table 6. Minority Status Subgroup Analysis: First Stage and Reduced Form Difference-in-Difference Estimates of the Effect of the Change in the Grading Scale for Minorities

VARIABLES	(1) First Stage	(2) Reduced Form
First Year STEM GPA> Non-STEM GPA	0.076	
	(0.023)**	
First Year GPA Diff.	0.156	
	(0.060)**	
First Year STEM GPA	0.264	
	(0.096)**	
First Year Non-STEM GPA	0.087	
	(0.102)	
STEM Major year 2		0.043
		(0.004)***
STEM Major year 3		0.058
		(0.010)***
Graduating in 6yrs with a STEM major		0.043
		(0.015)**
STEM credits attempted in six years or less		1.865
		(2.188)
Number of STEM courses enrolled in 6yrs or less		4.620
		(1.620)**
Non-STEM Major year 2		0.044
		(0.028)
Non-STEM Major year 3		0.013
		(0.029)
Graduating in 6yrs with a Non-STEM major		-0.018
		(0.080)
Non-STEM credits attempted in six years or less		6.626
		(5.449)
Number of Non-STEM courses enrolled in 6yrs or less		5.274
		(2.130)**
Did not earn a BA in 6yrs		-0.024
		(0.068)

Source: Author using student-transcript-level data from Florida Department of Education
Note: This table shows results for the first stage and reduced form estimates. Minorities include blacks non-Hispanics and Hispanics. These models include controls for covariates and institution fixed effects. Variables are imputed to zero if missing; missing data flags are included for all variables with missing data. Standard errors clustered at institution level. Sample size for the reduced form is 17,685. The covariates used these models include in dummy variables for female, US citizen, those who qualified for free lunch in grade 12, age and age squared at first entry, and STEM initial major intention, SAT reading and math scores, SAT reading and math scores squared. My sample is restricted to high school graduates who entered a 4-year public institution straight from high school for the first time in 1996, 2000 and 2003.

Table 7. Minority Status Subgroup Analysis: First Stage and Reduced Form Difference-in-Difference Estimates of the Effect of the Change in the Grading Scale for Non-Minorities

VARIABLES	(1) First Stage	(2) Reduced Form
First Year STEM GPA> Non-STEM GPA	0.074	
	(0.010)***	
First Year GPA Diff.	0.148	
	(0.036)***	
First Year STEM GPA	0.2893	
	(0.099)**	
First Year Non-STEM GPA	0.107	
	(0.067)	
STEM Major year 2		0.034
		(0.007)***
STEM Major year 3		0.038
		(0.008)***
Graduating in 6yrs with a STEM major		0.033
		(0.010)**
STEM credits attempted in six years or less		0.420
		(0.988)
Number of STEM courses enrolled in 6yrs or less		3.374
		(0.554)***
Non-STEM Major year 2		-0.023
		(0.011)*
Non-STEM Major year 3		0.005
		(0.025)
Graduating in 6yrs with a Non-STEM major		-0.020
		(0.060)
Non-STEM credits attempted in six years or less		2.637
		(4.153)
Number of Non-STEM courses enrolled in 6yrs or less		3.533
		(1.563)*
Did not earn a BA in 6yrs		-0.013
		(0.069)

Source: Author using student-transcript-level data from Florida Department of Education
Note: This table shows results for the first stage and reduced form estimates. Minorities include blacks non-Hispanics and Hispanics. These models include controls for covariates and institution fixed effects. Variables are imputed to zero if missing; missing data flags are included for all variables with missing data. Standard errors clustered at institution level. Sample size for the reduced form is 42,124. The covariates used these models include in dummy variables for female, US citizen, those who qualified for free lunch in grade 12, age and age squared at first entry, and STEM initial major intention, SAT reading and math scores, SAT reading and math scores squared. My sample is restricted to high school graduates who entered a 4-year public institution straight from high school for the first time in 1996, 2000 and 2003.

6. Discussion and Conclusion

Concerns over disparities in grading standards and grade inflation have led institutions to adopt various grading reforms. Changing the grading scale from whole-grade grading scale to plus/minus grading has been widely adopted as a way to curb grade inflation. Yet, the effect of this grading policy on STEM graduation and major choice has never been directly studied.

STEM majors are usually associated with higher and more rigorous grading standards than non-STEM majors. This policy was hypothesized to mitigate the effects of grade inflation by reducing grade differentials between STEM and non-STEM courses. By reducing grading disparities, students would be more attracted to STEM courses, which in turn may improve STEM graduation and major choice.

In the Fall semester of 2000 and 2001, two universities in Florida changed their grading scale. Using a difference-in-difference approach, I compare similar students over time at institutions with or without a change in grading scale so as to estimate the effect of this policy.

My analysis finds that this grading policy significantly decreases the first year GPA differential between non-STEM and STEM courses, by mainly increasing STEM grades. The consequences of changing the grading scale on STEM grades are theoretically ambiguous. My findings could be potentially explained by professors' preferences for "inflating" STEM grades to better differentiate students, particularly if relatively more students in STEM are at the margin of achieving higher grades than in non-STEM fields. Students who value higher grades could also respond to this policy by exerting more effort now that they can get better information regarding their potential for success or suitability for STEM courses.

I also find that after this grading policy students attending treated institutions were about 3 percentage points more likely to graduate in STEM in 6 years or less, which corresponds to an increase of about 40 percent. Finally, my analysis demonstrates that this grading policy has complex heterogeneous effects in the way students are discouraged from leaving STEM. This policy has larger effects for men than women and for racial minorities than non- minorities. However, for women the effects on STEM outcomes seem to be explained by a reduction in the probability of leaving STEM by switching into non-STEM, whereas for men these effects come from a decrease in the probability of dropping out of college. While women and men follow different STEM persistence patterns, the differences in STEM persistence patterns between racial minorities and non-minorities are not as affected by the change in the grading scale.

How these results are interpreted depends on whether they are viewed from an institutional or student perspective. From an institutional perspective, my results encourage the adoption of a plus/minus scale to improve STEM graduation outcomes and reduce drop-out rates. Yet, further research is needed to investigate long-term effects of this policy and why professors and/or students in STEM departments are more sensitive to this grading policy. Still, irrespective of the grading policy to be implemented at institutional level, this study sheds the light on the importance of understanding differential responses to grading policies among high-and low-grading departments that have been differentially affected by grade inflation. In addition to this, even within departments, students respond differently to grading policies depending upon where they are on the grade distribution and how much they value grades.

From the student perspective, the results indicate that students may be following higher grades, particularly women. If students do so as a response to grade inflation and grade differentials among departments and not because they develop new interests, then there is an argument for intervention. At first, it is not totally obvious that this policy improves the composition of STEM graduates because effects across subgroups are not significantly different.

But, the higher overall impact estimates for men are mainly explained by the fact that men are relatively more affected at the bottom of the grade distribution. Also, the effect on STEM graduation for women is fully explained by a decrease in the probability of switching into non-STEM, which suggests that for women grade differentials are perceived as a signal of relative ability in STEM versus non-STEM courses.

The existing literature of the effect of grading policies on major choices is scarce, and only one study has indirectly evaluated how reducing grade differentials between humanities (high-grading department) and economics departments (low-grading department) influence student's major choices at Wellesley (Butcher et al., 2014). My results contribute to this literature by demonstrating that students' major choices are affected by grades and that grading policies, intended to reduce grade inflation, increase graduation in STEM (low-grading departments). While the focus of this study is on two 4-year public institutions in Florida, these results might have broader applicability than those from the prior literature, which has been mainly focus on very selective institutions.[21] Moreover, this study evaluates the impact of a more sustainable and relatively cheaper grading policy that a growing number of institutions are adopting.

Finally, if we are interested in higher STEM graduation rates for racial minorities and women, before diverting more resources towards STEM departments by trying to make these departments more diverse, having more college guidance counselors, or even freezing STEM tuition rates, it may be more cost-effective to first adopt grading policies intended to reduce grade differentials so as to send reliable signals to students who are sensitive to grades.

[21] For example, while the national average SAT scores in reading and math between 1996 and 2003 are around 506 and 515, respectively, at these two Florida public institutions SAT scores are 538 and 544 for reading and math. In contrast, SAT scores at Wellesley are 605 and 611 for reading and math.

References

Altonji, Joseph G., Erica Blom, and Costas Meghir. (2012). *Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers*. Annual Review of Economics 4: 185-223

Altonji, Joseph G., Lisa B. Kahn, and Jamin D. Speer. (2014). *Trends in Earnings Differentials across College Majors and the Changing Task Composition of Jobs*. American Economic Review, American Economic Association, vol. 104(5): 387-393.

Anderson, E. L., & Kim, D. (2006). *Increasing the success of minority students in science and technology*. American Council on Education Washington, DC.

Arcidiacono, Peter, Esteban Aucejo, Hanming Fang, and Kenneth Spenner. (2011). *What Happens After Enrollment? An Analysis of the Time Path of Racial Difference in GPA and Major Choice*. Quantitative Economics 2(3), 303-333.

Arcidiacono, P., Aucejo, E. M., and 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(1), 1–24.

Attewell, P., Heil, S., and Reisel, L. (2011). *Competing Explanations of Undergraduate Noncompletion*. American Educational Research Journal, 48(3), 536–559.

Baker III, H. E., and Bates, H. L. (1999). *Student and faculty perceptions of the impact of plus/minus grading: a management department perspective*. Journal on Excellence in College Teaching, 10(1), 23–33.

Barnes, K. D., and Buring, S. M. (2012). *The Effect of Various Grading Scales on Student Grade Point Averages*. American Journal of Pharmaceutical Education, 76(3).

Barnes, L. L. B., Bull, K. S., Campbell, N. J., and Perry, K. M. (2001). *Effects of Academic Discipline and Teaching Goals in Predicting Grading Beliefs Among Undergraduate Teaching Faculty*. Research in Higher Education, 42(4), 455–467.

Betts, J. R. and J. Grogger. (2003). *The impact of grading standards on student achievement, educational attainment, and entry-level earnings*. Economics of Education Review 22: 343–352.

Bressette, A. (2002). *Arguments for Plus/minus Grading: A Case Stud*y. Educational Research Quarterly, 25(3), 29.

Brown, C., and Corcoran, M. (1996). *Sex-Based Differences in School Content and the Male/Female Wage Gap* (Working Paper No. 5580). National Bureau of Economic Research.

Brumfield, Cody. (2005). *Current Trends in Grades and Grading Practices in Higher Education: Results of the 2004 AACRAO Survey.* Washington, D.C.: American Association of Collegiate Registrars and Admissions Officers (AACRAO).

Butcher, K. F., McEwan, P. J., and Weerapana, A. (2014). *The Effects of an Anti-Grade-Inflation Policy at Wellesley College*. The Journal of Economic Perspectives, 28(3), 189–204.

Chen, X. (2013). *STEM Attrition: College Students' Paths into and out of STEM Fields.* Statistical Analysis Report. NCES 2014-001. National Center for Education Statistics.

Chizmar, J. F. (2000). *A Discrete-Time Hazard Analysis of the Role of Gender in Persistence in the Economics Major*. The Journal of Economic Education, 31(2), 107–118.

Crisp, G., Nora, A., and Taggart, A. (2009). *Student Characteristics, Pre-College, College, and Environmental Factors as Predictors of Majoring in and Earning a STEM Degree: An Analysis of Students Attending a Hispanic Serving Institution*. American Educational Research Journal, 46(4), 924–942.

DeBerard, M. S., Spielmans, G. I., and Julka, D. L. (2004). *Predictors of Academic Achievement and Retention among College Freshmen: A Longitudinal Study*. College Student Journal, 38(1), 66.

Dynan, K. E., and Rouse, C. E. (1997). *The Underrepresentation of Women in Economics: A Study of Undergraduate Economics Students*. The Journal of Economic Education, 28(4), 350–368.

Farland, R., and Cepeda, R. (1989). *Plus and minus grading.* (Report No. ED 309 802). Sacramento: Prepared as Agenda Item Number 9 at a meeting of the Board of Governors of the California Community Colleges, September 14-15. (ERIC Document Reproduction Service No. JC 890 370).

Figlio, David N. and M. E. Lucas. (2004). *Do high grading standards affect student performance?* Journal of Public Economics 88 _9–10_: 1815–1834.

Florida Board of Governors State University System. *Degrees Awarded by State University System Institutions since 1991*. Retrieved from http://www.flbog.edu/resources/iud/degrees_results.php

Gerber, T. P., and Cheung, S. Y. (2008). *Horizontal Stratification in Postsecondary Education: Forms, Explanations, and Implications*. Annual Review of Sociology, 34(1), 299–318.

Grant, D., and W. Green. (2013). *Grades as incentives*. Empirical Economics 44(3): 1563–92.

Grove, W. A., and T. Wasserman. (2006). *Incentives and student learning: A natural experiment with economics problem sets*. Research on Teaching Innovations 96(2): 447–52.

Griffith, A. L. (2010). *Persistence of women and minorities in STEM field majors: Is it the school that matters?* Economics of Education Review, 29(6), 911–922.

Hill, C., Corbett, C., and St. Rose, A. (2010). *Why So Few? Women in Science, Technology, Engineering, and Mathematics*. American Association of University Women.

Huang, G., Taddese, N., and Walter, E. (2000). *Entry and Persistence of Women and Minorities in College Science and Engineering Education*. Education Statistics Quarterly, 2(3), 59–60.

Hurtado, S., Eagan, K., and Chang, M. (2010). *Degrees of success: Bachelor's degree completion rates among initial stem majors*. Higher Education Research Institute at UCLA, January.

Jacobson, Louis and Christine Mokher (2009). *Pathways to Boosting the Earnings of Low-Income Students by Increasing Their Educational Attainment*. Prepared for the Bill & Melinda Gates Foundation by the Hudson Institute Center for Employment Policy and CNA Analysis & Solutions, January 2009.

Jensen, E. J., and Owen, A. L. (2001). *Pedagogy, Gender, and Interest in Economics*. The Journal of Economic Education, 32(4), 323–343.

Johnson, V. E. (2003). *Grade Inflation: A Crisis in College Education*. Springer Science & Business Media.

Kokkelenberg, E. C., & Sinha, E. (2010). *Who succeeds in STEM studies? An analysis of Binghamton University undergraduate students*. Economics of Education Review, 29(6), 935–946.

Nora, A., Cabrera, A., Hagedorn, L. S., and Pascarella, E. (1996). *Differential impacts of academic and social experiences on college-related behavioral outcomes across different ethnic and gender groups at four-year institutions*. Research in Higher Education, 37(4), 427–451.

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

Quann, C. J. (1987). *Plus-Minus Grading: A Case Study and National Implications* (pp. 1-17). Washington, D.C.: American Association of Collegiate Registrars and Admissions Officers.

Rask, K. (2010). *Attrition in STEM fields at a liberal arts college: The importance of grades and pre-collegiate preferences*. Economics of Education Review, 29(6), 892–900.

Rask, K. N., & Bailey, E. M. (2002). *Are Faculty Role Models? Evidence from Major Choice in an Undergraduate Institution*. The Journal of Economic Education, 33(2), 99–124.

Rask, K., & Tiefenthaler, J. (2008). *The role of grade sensitivity in explaining the gender imbalance in undergraduate economics*. Economics of Education Review, 27(6), 676–687.

Riley, H. J., Checca, R.C., Singer, T.S., Worthington, D.F. (1994). *Current Trends in Grades and Grading Practices in Undergraduate Higher Education*. The Results of the 1992 AACRAO Survey. New York: 1-76.

Robb, R. E., and Robb, A. L. (1999). *Gender and the Study of Economics: The Role of Gender of the Instructor.* The Journal of Economic Education, 30(1), 3–19.

Rojstaczer, S., & Healy, C. (2010, March 4). *Grading in American Colleges and Universities*. Retrieved November 26, 2014, from http://www.gradeinflation.com/tcr2010grading.pdf

Sabot, R., & Wakeman-Linn, J. (1991). *Grade Inflation and Course Choice*. The Journal of Economic Perspectives, 5(1), 159–170.

Seymour, E., and Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences* (Vol. 12). Westview Press Boulder, CO.

Shannon, J. L. (1979). *High School Grades--How Fair Are Plus and Minus Suffixes?* Education, 100(2), 153–57.

Shaw, E. J., and Barbuti, S. (2010). *Patterns of Persistence in Intended College Major with a Focus on STEM Majors*. NACADA Journal, 30(2), 19–34.

Stinebrickner, T. R., and Stinebrickner, R. (2009). Learning about Academic Ability and the College Drop-out Decision (Working Paper No. 14810). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w14810

Stinebrickner, T. R., and Stinebrickner, R. (2011). *Math or Science? Using Longitudinal Expectations Data to Examine the Process of Choosing a College Major* (Working Paper No. 16869). National Bureau of Economic Research.

Strenta, A. C., Elliott, R., Adair, R., Matier, M., and Scott, J. (1994). *Choosing and leaving science in highly selective institutions*. Research in Higher Education, 35(5), 513–547.

Thomas, G. E. (1985). *College Major and Career Inequality: Implications for Black Students.* The Journal of Negro Education, 54(4), 537–547.

U.S. Department of Homeland Security, Immigration and Customs Enforcement. *STEM Designated Degree Programs*. Retrieved April 30, 2010, from http://www.ice.gov/sevis/stemlist.htm.