The Impact of Hidden Grades on Student Decision-Making and Academic Performance: An Examination of a Policy Change at MIT

Abstract

Colleges and universities work hard to create environments that encourage student learning, and they develop grading policies, in part, to motivate their students to perform well. Grades provide two kinds of information about a student's abilities and learned knowledge: *internal* information that informs the students themselves about the university's assessment of their talents and competencies; and *external* information that informs faculty, other institutions, and potential employers about student performance.

At the Massachusetts Institute of Technology (MIT), freshman grading policies were changed in the fall of 2002 in an effort to better prepare freshmen for the academic rigors of sophomore year and beyond. Prior to the 2002-03 academic year, all freshmen at MIT received "hidden" grades in *both* semesters of their freshman year. A hidden grade is a *letter* grade that is communicated to the student but is recorded as *pass/no-record* on the student's official transcript.

Beginning in the fall of 2002, freshmen received hidden grades for the *first semester only* of their freshman year. Therefore, pre- and post-2002 freshmen received the *same* internal information on their grades in the second semester, but post-2002 freshmen were subject to this information being shared externally. In this study, I estimated the causal impact of MIT's having *hidden* versus *externally-shared* grades on subsequent student decision-making and academic

performance by taking advantage of the natural experiment that was inaugurated by this policy change.

I looked specifically at the impact of the grading-policy change on freshman spring semester GPA, credit units taken, the probability of declaring early sophomore status, and the probability of taking a more mathematically advanced version of Physics II. I found that freshmen with *externally-shared* grades, on average, earned higher GPAs, had a higher probability of declaring early sophomore standing, took slightly fewer credit hours, and had a slightly lower probability of taking a more rigorous version of Physics II, compared to freshmen with *hidden* grades second semester. Also, for three of my four outcomes, I found that the estimated effect of the grading-policy change differed by the level of a student's pre-college academic performance.

Table of Contents

Introduction: Background and Context of the Study

Colleges and universities work hard to create environments that encourage student learning, and they develop grading policies, in part, to motivate their students to perform well (Guskey and Bailey 2001, Michaels 1977). At most institutions of higher education, evaluation of student performance is often formalized through the use of grades that appear on a student's official academic transcript (Olsen, 1975).[1] Grades provide two kinds of information about a student's abilities and learned knowledge: *internal* information that informs the students themselves about the university's assessment of their talents and competencies; and *external* information that informs faculty, other institutions, and potential employers about student performance (Cherry and Ellis, 2005). At the Massachusetts Institute of Technology (MIT), freshman grading policies were changed in the fall of 2002 in an effort to better prepare freshmen for the academic rigors of sophomore year and beyond (MIT 2008).

Prior to the 2002-03 academic year, all freshmen at MIT received "hidden" grades in *both* semesters of their freshman year. A hidden grade is a *letter* grade that is communicated to the student but is recorded as *pass/no-record* on the student's official transcript.[2] Under this system, students know their grades, but these grades do not appear on the external transcript and therefore communicate no message to other faculty, institutions, and future employers. In other words, students receive *internal* information about their performance, but this information is not shared with *external* parties. However, beginning in the fall of 2002, freshmen received hidden

[1] These grades can be broadly categorized into two types: (a) traditional letter grades, and (b) pass-fail grades. In a traditional grading system, students receive scaled letter grades, each usually associated with a numeric value – i.e. A=4; B=3; C=2; D=1; F=0 – with the grades that fall earlier in the alphabet, and have higher numeric values, corresponding to assessments of higher academic quality for the work. In a pass-fail grading system, students receive a "pass" if they meet a defined minimum level of performance and a "fail" if they do not (Guskey and Bailey, 2001).

[2] Hidden grades do not contribute to a student's overall GPA.

grades for the *first semester only* of their freshman year. Therefore, pre- and post-2002 freshmen received the *same* internal information on their grades in the second semester, but post-2002 freshmen were subject to this information being shared externally. In this study, I estimated the causal impact of MIT's having *hidden* versus *externally-shared* grades on subsequent student decision-making and academic performance by taking advantage of the natural experiment that was inaugurated by this policy change.

Figure 1 illustrates the type of grading system freshmen faced before and after the policy change. In the fall semester, pre- and post-policy, freshmen were under a *pass/no-record* system. Under this system, freshmen received a *pass* on their official transcript for each course they took in which they received a grade of C or better. If they received a grade of D or F in a course, there was no record of this on the transcript. In the spring semester, pre-policy freshmen remained under a *pass/no-record* system. Post-policy freshmen, however, were graded under an *A/B/C/no-record* system. Under this system, grades of A through C appeared on the external transcript, but not the "non-passing" grades of D through F. There was no indication on the student transcript that the student took a course for which he or she received a non-passing grade. In other words, in the second semester after the policy shift, grades were *unhidden* on the pass side (A,B,C) and *no-record* on the fail side (D or F). Therefore, grades were only partially unhidden in the spring semester during the post-policy period.[3]

The rationale for providing hidden grades at MIT, which began in 1982, was to relieve freshman anxiety and pressure during the year of transition from high school to college, to encourage acquisition of foundational knowledge, to encourage the development of social skills,

[3] Pre- and post-policy sophomores (and beyond) were graded using non-hidden letter grades. The change from pass/no-record to A/B/C/no-record could be considered a gradual transition from fully hidden grades to fully non-hidden grades.

to give freshmen a sense of freedom to make a wider choice of courses (course exploration), and to compensate for differences in secondary school preparation (MIT 2000). Similar reasons motivated other colleges and universities in the 1960s and 1970s to supplement their traditional grading systems with pass-fail alternatives (Olsen 1975). However, in 2002, MIT decided to eliminate hidden grades for freshmen in their *second* semester because some faculty felt that the previous policy encouraged complacency and interfered with students' academic performance in the freshman year and beyond. These beliefs are consistent with the broader research literature on pass-fail grading, which indicates that when students elect the pass-fail option they earn lower grades (had those grades been assigned) than those who receive traditional letter grades (Giometti 1976, Olsen 1975, Sgan 1970, Stallings and Smock 1971, Wittich 1972), with freshmen experiencing a larger drop in GPA than upperclassmen (Olsen 1975, Sgan 1970). The literature also suggests that students who take pass-fail courses tend not to explore courses outside of their major (Olsen 1975, Stallings and Smock 1971, Wittich 1972). These differences, however, are potentially *endogenous* as they typically involve the selection of grading options by students: students who think that they will not do well may be less likely to choose the letter grade option. Moreover, the literature on grading options is several decades old and may not reflect current conditions or effects. In my study, I addressed the concerns about endogeneity – and extended the literature – by providing estimates of the causal impact of grading policies on subsequent student decision-making and academic performance.

At MIT, I assume that the grading-policy change that I have referred to – the elimination of hidden grades in the second semester – occurred *exogenously*. That is, the assignment of freshmen to a particular form of the grading system was determined solely by the policy shift (pre-policy freshmen received hidden grades in both semesters of freshman year; post-policy freshmen received hidden grades in only the first semester of freshman year) and not student choice. Consequently, with no change in admissions policies, I argue that freshmen were *equal in expectation* – indistinguishable on all observed and unobserved characteristics, on average, in the population – prior to, and after, the policy change.[4] This change in grading policy then provided me with a natural experiment with which to investigate the causal impact of having unhidden grades on subsequent student decision-making and academic performance.

<u>The Potential Impact of Hidden Grades</u>

My research questions were driven by how students might react to a discrete grading-policy change. Theoretically, introducing the external signal of grades in the second semester of the freshman year (i.e., eliminating hidden grades) could have several possible effects:

1. <u>Change the level of effort in courses</u>. Students may not take their courses as seriously under a system of hidden grades, compared to a system of non-hidden grades. As a result, students may perform worse academically when grades are hidden from the transcript, as has been found in the previous literature. Intuitively, freshmen may work harder in their classes when they know that the grades that they receive will be recorded and shared with external parties. In other words, hidden grades may act as a disincentive for students to take their coursework seriously.

From my previous work, using a sample of two cohorts of students at MIT, I found that eliminating hidden grades led to higher freshman spring semester GPA (Harris 2009), as theory would predict. I also found that this positive effect on GPA did not persist to sophomore year. This is perhaps because students both pre- and post- policy change had externally-shared grades

[4] The validity of this assumption is explored in the analysis by comparing observable characteristics of students before and after the policy. Additional threats to the validity of this assumption are also discussed later in the paper.

during their sophomore year. My previous results are bolstered by an internal MIT descriptive study that examined freshman spring term GPA performance before and after the grading-policy change. The study found that the percentage of D and F grades (of all grades given over courses and students) decreased as a result of the change to A/B/C/no-record grading, from six percent pre-policy to four percent post-policy (MIT 2008).[5] My study investigated this question more rigorously using causal inference and data on eight cohorts of students.

2. Change timing of declaring sophomore standing. Students may be less likely to opt into sophomore standing under a system of hidden grades, given that sophomores are then graded using non-hidden letter grades. Before spring semester, freshmen are notified if they are eligible for early sophomore standing. Eligibility is based mostly on the percent completion of university requirements, namely they must have completed at least one-quarter of the undergraduate program and taken specific core courses.[6] The primary benefit of being an early sophomore is that a student can declare a departmental major and be assigned a faculty advisor in that major. If a student declares sophomore standing, however, that student is then graded under the sophomore grading system, with fully non-hidden grades.

Intuitively, eligible freshmen may be more inclined to declare early sophomore standing post-policy than pre-policy, because the post-policy freshman spring term grading system resembles the sophomore grading system more closely than the pre-policy freshman grading system. In the pre-policy period, students who entertained declaring early sophomore standing faced the choice between *pass/no-record* (as a freshman) and fully *non-hidden grades* (as a

[5] Rather than focusing on failing grades, my study looked at the impact of the grading-policy change on average GPA.
[6] Students are eligible for early sophomore standing if they have completed one-fourth of the credits required for earning a degree by the end of their first semester, including a communication intensive course and a majority of the science requirement courses (mathematics, physics, chemistry, and biology). Transfer credits are included in this count.

sophomore). In the post-policy period, the choice was between *A/B/C/no-record* (as a freshman) and fully *non-hidden grades* (as a sophomore). So, while there was some benefit to maintaining freshman standing status (namely not being penalized for earning a very low grade in a course), the incentive to remain a freshman diminished greatly with the grading-policy change. Indeed, based on an internal descriptive review of MIT's freshman grading system, the number of early sophomores increased substantially in the post-policy period, compared to the pre-policy period (MIT 2008).

3. <u>Alter the number of credits taken</u>.[7] Undergraduates must complete a certain number of credit units in order to graduate. Students may choose to take more credit units during semesters with no external signal for grades, given that there is no penalty on the student transcript for earning a lower grade (e.g., C) under the hidden grade system. In other words, post-2002 freshmen may *reduce* the number of credit units taken spring semester freshman year to allow for more time to devote to their coursework. This hypothesis is consistent with a finding from a 2008 review by MIT of the freshman grading-policy change using purely descriptive analyses. A subcommittee of the Committee on the Undergraduate Program showed that the total number of courses taken spring term by freshmen dropped once the new grading policy was implemented (MIT 2008). In my study, I focused on credit units as an outcome rather than the number of courses. This is because different courses carry different credit loads, and credit units may therefore provide a better approximation of time spent on academic coursework than the number of courses taken.

[7] Credit units are defined as the total number of hours spent each week in class/laboratory for a given course plus the estimated time that the average student spends each week in outside preparation for that course. At MIT, a full-time freshman must register for at least 36 credit units each semester and may not register for courses totaling more than 54 credit units in the fall term and 57 credit units in the spring term.

4. Take different courses. Freshmen may opt to take, or not to take, particular courses, depending on the grading system. For example, if students are given a choice between two courses of similar content, one of which is perceived to be more difficult than the other, students may be less likely to take the more difficult course if the grade they receive shows up on their official transcript. To test this theory, I focused on a sequence of Physics courses that are part of MIT's *General Institute Requirement* (GIR).[8] At MIT, every student must take, or test out of, two courses in physics: Physics I and Physics II. Most students take Physics II in the spring semester of their freshman year. In fact, in my sample, 74% of students took Physics II in the Spring.

MIT offers different versions of Physics II, but they can be roughly categorized into a "standard" version and a "more advanced" version. Both versions fulfill the same course requirement.[9] In other words, it is left up to the student which version he or she elects to take, regardless of major. The more advanced version is described in course materials as being more "rigorous" and more "mathematically advanced" than the standard version. No doubt students who enter MIT with interests in Physics and Mathematics are more likely to take the more advanced version of Physics II, compared to students with other pre-college major interests, but the choice is left up to the student.[10] This voluntary feature lets me test my hypothesis that

[8] MIT's GIR consists of 17 courses: 9 in math, science, and technology and 8 in the humanities, arts, and social sciences.
[9] The two versions of Physics II share the same pre-requisites. The more advanced version also has a co-requisite of Calculus II. Calculus II is a required course for all MIT students, but students who take the standard version of Physics II can opt to take Calculus II in a subsequent semester.
[10] The more advanced Physics II course did not substantively change in terms of content or pedagogy in the years of my sample. The standard Physics II course, however, did. Specifically, the teaching format changed from lecture/recitation to a more collaborative, technology-driven format in the same year as the grading policy change. In addition, in the early years of my sample, some students took an experimental version of the standard Physics II course using a hands-on/lab-based format. A potential threat to validity is that some students who would have taken the standard Physics II course, but did not like interactive learning pedagogies, may have chosen to take the more advanced Physics II course instead.

students are less likely to opt into courses that are perceived to be more difficult if the grades in those courses are unhidden.

In addition to the four potential responses to the grading-policy change described above, students may also react in other ways. For example, it may be that students are less likely to explore courses outside of their comfort zone under a non-hidden grading policy because students may perceive that the risk of getting a low grade is higher in unfamiliar courses. As stated earlier, one of the original motivations for MIT's adoption of a hidden grading system was to allow students the freedom to explore new areas of study. The literature, however, suggests that students who take pass-fail courses tend not to venture far from their intended major (Olsen 1975, Stallings and Smock 1971, Wittich 1972). I did not test this outcome in my study, in part, because my research was limited to a student's first year in college, during which they had limited opportunities to take elective courses. In addition, I had no appropriate proxy for course exploration. Answering this question is certainly a worthy endeavor, though, and would be a natural extension of my current study.

Alternatively, students may choose to take the exact same courses during college but alter the timing of those courses, perhaps in an effort to bolster their GPA. For example, they may take courses perceived to be more difficult during fall term when they have hidden grades and courses perceived to be easier during subsequent terms when grades are no longer hidden. Similar to the discussion on course exploration, I did not address this question in my study.

With all four of my outcomes, it is possible that decision-making and academic performance at MIT remained unchanged after inauguration of the new grading system, if students, on average, were motivated to act based largely on the *internal* information they received about their grades rather than the *external* signal of the letter grades. Also, it is possible that students do not care about, nor react to, the grades that they receive. If this is the case, then students may not change any of their behaviors in response to a grading-policy change. However, there are ample reasons to believe that both internal and external information on grades do matter to at least some students. Regarding *internal* information, the literature suggests strongly that student course-taking decisions are affected by the grades students receive in prior courses, and that students are less likely to pursue a path of study if they earn poor grades in introductory courses in that subject (Arcidiacono 2003, Montmarquette et. al. 2002, Rask and Bailey 2002, Rask and Tiefenthlaer 2004, Sabot and Wakeman-Linn 1991). Regarding *external* information, three-quarters of undergraduate students at MIT eventually attend graduate school (MIT 2005). College transcripts are an important component of admission into such schools. In addition, some employers request transcripts during the hiring process (conversation with MIT Careers Office, July 2009). Finally, students who attend MIT and other selective institutions come from an environment in which grades are highly valued.

<u>Research Questions</u>

For my study, I examined the causal impact of no longer having hidden grades on subsequent student decision-making and academic performance. To investigate the four potential effects on student behavior described above, my specific research questions were:

1. After controlling for selected observable characteristics, did freshman students *earn similar grades* during their second semester after MIT went from a policy in which grades were externally hidden for both semesters of freshman year to having only the first semester grades of freshman year hidden?

2. After controlling for selected observable characteristics, did a similar percentage of freshman students *declare early sophomore standing* during their second semester after MIT went from a policy in which grades were externally hidden for both semesters of freshman year to having only the first semester grades of freshman year hidden?
3. After controlling for selected observable characteristics, did freshman students *enroll in a similar number of credit units* during their second semester after MIT went from a policy in which grades were externally hidden for both semesters of freshman year to having only the first semester grades of freshman year hidden?
4. After controlling for selected observable characteristics, did a similar percentage of freshman students *take a more mathematically advanced Physics II course* during their second semester after MIT went from a policy in which grades were externally hidden for both semesters of freshman year to having only the first semester grades of freshman year hidden?

Research Design

Site

My study was of MIT, a selective, coeducational, private research university located in the Northeast of the United States. MIT is an appropriate site for this study because, in 2002, there was a shift in the freshman grading policy, which I argue was exogenous, thus enabling me to draw unbiased causal inferences about the impact of the policy change on subsequent student decisions and academic performance. By exogenous, I mean that I assume that the grading-policy change had no effect on the college choices of potential MIT students and no effect on

MIT faculty behavior.[11] MIT has a total student population of approximately 10,000 students, 40% of whom are undergraduates. Of the undergraduate population, 45% are women. Over 80% of undergraduates at MIT major in science and engineering.

While MIT is obviously a unique institution given its focus on science and engineering, its student body has similar characteristics to other highly selective colleges and universities.[12] As such, the findings of this study may have some external validity for these types of institutions. The generalizability, however, may be limited for other types of schools. Moreover, it is important to recognize that MIT's freshman grading policy is somewhat different than those at other institutions described in the literature. At MIT, freshmen received the same *internal* information about their grades pre- and post-policy change. In the literature, it is uncommon for students who elect pass-fail grades to know officially what grade they would have received had they not chosen this option (Giometti 1976, Sgan 1970, Stallings and Smock 1971, Wittich 1972).

Sample

My sample included the eight cohorts of undergraduate students who entered MIT as first-time freshmen from Fall of 1998 to Fall of 2005, 8,098 students in total.[13] I had almost no missing data on any of my variables. I dropped just 16 cases (less than 0.2% of my original sample) due to missing admissions data. Of the 16 cases, only five records had no admissions

[11] I discuss potential threats to the validity of my assumptions later in the paper.
[12] MIT is a member of the Consortium on Financing Higher Education (COFHE), a group of thirty-one selective, private colleges and universities. Students at these schools are highly-talented and motivated and tend to have stellar high school records and standardized test scores.
[13] For my fourth research question, I limited my sample to students who took a version of Physics II in the spring semester. For this research question, my sample was 6,018 students, or 74% of my total sample.

data whatsoever. I found no systematic reason for these missing cases, and so I am not concerned that their elimination impacted my results in any way.

The first four cohorts of these freshmen entered MIT under the pre-policy grading system (hidden grades both semesters freshman year); the second four cohorts of freshmen entered MIT under the post-policy grading system (hidden grades first semester only). With this sample size, at usual levels of *Type I* error, I had a statistical power of at least .80 for detecting small effects (δ=0.10) (Light, Singer, and Willett 1990).

Measures

My data included record-level admissions, financial aid, and transcript information on each student in my sample for each student's freshman year from MIT administrative databases. I had data on every course each student took, their grades in those courses, their pre-admissions test scores, and pre-college preference for major. I review the variables that I included in my analyses below briefly, and I provide additional details in Table 1.

Outcomes: I defined the following variables for the second semester of freshman year, the first term during which students before and after the policy change faced the different grading policies.

1. $SPRINGGPA_i$ is a continuous variable that measures a student's grade point average in the spring semester. The values range from zero to five, with higher values representing a higher grade point average (A=5, B=4, C=3, D=2, F=0).

2. $SOPHOMORE_i$ is a dichotomous variable that indicates whether or not the student declared early sophomore standing in the spring semester (1=early sophomore standing; 0=freshman standing).

3. $CREDITS_i$ is a continuous variable that records the total number of credit units taken by a student in the spring semester.[14]

4. $DIFFICULT_i$ is a dichotomous variable that indicates whether or not the student took a more mathematically advanced Physics II course in the spring semester (1=took more mathematically advanced Physics II course; 0=took standard Physics II course or equivalent).

Question Predictor: For all four research questions, my key independent variable is NOT_HIDDEN_i, a dichotomous variable that indicates whether or not the student received hidden grades during second semester of freshman year. Students who attended MIT before the policy change had a value of zero while students after the policy change had a value of one.

Control Variables: I included the following covariates in my statistical models: (a) student demographic variables (gender, race/ethnicity, citizenship, and financial need), (b) student pre-college major preference, and (c) student pre-college admissions scores. These variables have been found to be important in previous studies of grade performance and course selection (Arcidiacono 2003, Harris 2009, Montmarquette et. al. 2002, Rask and Bailey 2002, Rask and Tiefenthaler 2004, Sabot and Wakeman-Linn 1991). One of my principal control variables is SCORE_OBJ, which measures a student's pre-college academic performance. This

[14] While freshmen at MIT face a 57 credit limit maximum in the spring semester, exceptions are made if a freshman participates in a freshman seminar, whereby freshmen are allowed to exceed the cap by up to six units, depending on the seminar. In my sample, there were 288 freshmen who exceeded the 57 credit unit cap (less than four percent of my entire sample). Less than one percent of my sample exceeded the limit by five or more credit units.

numeric index ranges from 1 to 5 and is comprised of standardized test scores (e.g., SAT, ACT, TOEFL) relating to math, science, and humanities, as well as high school GPA and class rank. I expected that SCORE_OBJ would be positively related to my outcomes because previous studies on the college academic experience point to the combined predictive power of pre-college standardized test scores and high school grades on college success (Camara and Echternacht 2000). Additionally, I tested for an interaction effect between this variable and NOT_HIDDEN, my variable of interest, because I hypothesized that there may have been a differential effect of the grading-policy change on my outcomes based on the level of a student's pre-college academic performance. For example, it is conceivable that students with lower prior preparation (who are likely to have lower performance in college) would be more worried about having their grades visible on the transcript and react more strongly to the grading-policy change compared to students with higher prior preparation. Finally, I included a covariate for the year the student entered MIT ($YEAR_i$), centered on the year in which the grading policy changed, to control for secular (linear) trends in the outcome by year.[15]

<u>Data-Analysis</u>

I addressed three of my four research questions using the same sample. For my fourth question, concerning the outcome DIFFICULT, I limited my sample to students who enrolled in a version of Physics II in their spring semester. For my continuous outcomes (SPRINGGPA and CREDITS), I used ordinary least-squares (OLS) regression analysis. For my dichotomous research questions (SOPHOMORE and DIFFICULT), I used logistic regression analysis.

[15] By including this variable I had, in effect, an *interrupted time series* (ITS) design.

For all four outcomes, I proceeded with my regression modeling in a similar fashion. First, I specified and fitted a baseline model that contained my variable of interest, NOT_HIDDEN, and the linear effect of the year that the student entered MIT, centered on the year the grading policy changed (YEAR). This baseline model gives me the estimated relationship between my outcome and NOT_HIDDEN, controlling for year-to-year trends. Then, in my second specification, I added controls for selected student characteristics (gender, race/ethnicity, citizenship, financial need). In my third specification, I added controls for student pre-college academic standing (objective and subjective) and pre-college major interests.[16] The addition of these controls account for possible effects due to changing student body characteristics over time. Given my assumption that the grading-policy change occurred exogenously, adding these controls should not alter the baseline estimate. Finally, I added the interaction term NOT_HIDDEN * SCORE_OBJ to test whether the impact of the grading-policy change differed by a student's pre-college academic performance. To avoid redundancy, I describe here how I analyzed continuous outcome SPRINGGPA using OLS regression methods and dichotomous outcome SOPHOMORE using logistic regression methods.[17] The other two outcomes were analyzed in the same way, depending on their distributional properties.

To answer research question 1, I examined if after controlling for selected observable characteristics freshman students *earned similar grades* during their second semester after MIT went from a policy in which grades were externally hidden for both semesters of freshman year

[16] For outcome variable DIFFICULT, I used department level controls for interest in Physics and Math rather than the more broadly defined school-level interest variables. I did this because, intuitively, students who took a more mathematically advanced Physics II course were probably influenced, to some degree, by their interest in Physics and Math specifically.

[17] I had eight years of data in my sample, but the grading policy disruption occurred in only one year. To make sure that I modeled the relationship between my outcome and predictors appropriately on either side of the disruption (i.e., estimate the effect of the treatment unbiasedly at the cut-point), I tested for the two-way interaction NOT_HIDDEN * YEAR. For each of my research questions, the interaction was not statistically significant, so I did not include the interaction term in my regression models.

to having only the first semester grades of freshman year hidden. I fitted the following multivariate regression model using OLS methods:

$$(1)\quad SPRINGGPA_i = \beta_0 + \beta_1 NOT\_HIDDEN_i + \beta_2 YEAR_i + \kappa' \mathrm{X}_i + \gamma' \mathrm{Z}_i + \varepsilon_i$$

where X represents a vector of student characteristics accompanied by regression parameters κ, Z represents a vector of pre-college ability/interest covariates accompanied by regression parameters γ, and ε_i is a student-level residual. Parameter β_1 represents the effect of the freshman grading-policy change, from having hidden grades to making them externally visible, on freshman second semester GPA. If an estimate of this parameter is statistically significant and positive, I conclude that the shift in grading policy was associated with a higher freshman spring semester GPA, controlling for student demographics and pre-college ability/interest. Parameter β_2 tells me whether or not I have a linear secular trend.[18] To examine whether the impact of the policy change differed by student pre-college achievement, I also tested to see if there was a statistically significant two-way interaction between NOT_HIDDEN and the level of a student's pre-college academic performance (SCORE_OBJ) on SPRINGGPA. I addressed my third research question regarding spring semester credit units, CREDITS, in a similar fashion as for SPRINGGPA.

To answer research question 2, I examined if after controlling for selected observable characteristics a similar percentage of freshman students *declared early sophomore standing* during their second semester after MIT went from a policy in which grades were externally hidden for both semesters of freshman year to having only the first semester grades of freshman year hidden. To address this research question, I fitted the following logistic regression model:

[18] If the coefficient is zero, I conclude that there is no trend. If there is a trend, it is important for me to model it, as this will increase my power by reducing residual variance.

$$(2)\quad \log_e\left[\frac{prob(SOPHOMORE_i=1)}{prob(SOPHOMORE_i=0)}\right]=\alpha_0+\alpha_1 NOT\_HIDDEN_i+\alpha_2 YEAR_i+\kappa'\mathrm{X}_i+\gamma'\mathrm{Z}_i$$

where X represents a vector of student demographic characteristics accompanied by regression parameters κ, and Z represents a vector of pre-college ability/interest covariates accompanied by regression parameters γ. Parameter α_1 represents the effect of the freshman grading-policy change, from having hidden grades to making them externally visible, on the log-odds of a student declaring sophomore standing second semester freshman year.[19] If an estimate of this parameter is statistically significant and positive, then I conclude that the shift in grading policy was associated with a higher probability of becoming an early sophomore, controlling for covariates. To test whether there was a differential effect by student pre-college performance, I also added a two-way interaction between NOT_HIDDEN and the level of a student's pre-college academic performance (SCORE_OBJ) on SOPHOMORE. I addressed my fourth research question regarding taking a more rigorous Physics II course in the spring semester, DIFFICULT, in a similar fashion as for SOPHOMORE.

Findings

My analytic approach rested on the assumption that the grading-policy change was exogenous. Unlike the previous literature, I argue that my study avoided comparisons of students who were under different grading options due to endogenous choices. Past studies have been unable to support unbiased causal conclusions as they did not account for unobserved influences on the grade choice, such as a student's own expectations of performance. In the case of this study, the assignment of freshmen to a particular form of the grading system was determined

[19] Interpreting changes in the log-odds is not very intuitive, so I also computed an odds-ratio by taking the natural log of my NOT_HIDDEN coefficient.

solely by the policy shift and not student choice. Therefore, I argue that freshmen before and after the grading-policy change were equal in expectation and identical, on average, in the population.

While I assume that the sudden change in grading policy at MIT was exogenous, if it did lead to changes in the types of students who subsequently enrolled at MIT, then my results may be biased. However, based on observable student characteristics, pre-policy freshmen looked very similar to post-policy freshmen. In Table 2, I present descriptive statistics of selected covariates pre- and post-policy change. Notice that freshmen who entered MIT during the pre-policy period (1998-2001) looked very similar to freshmen who entered MIT during the post-policy period (2002-2005). For example, the gender, race/ethnicity, and citizenship makeup of freshmen, pre-policy versus post-policy, differed by two percentage points or less.[20] Year-to-year incoming student demographics fluctuate very little at MIT, and based on Table 2, I have no reason to suspect that the profile of the freshman class pre- and post-policy period changed as a result of the grading-policy change. In fact, students conceivably did not even know about the policy change prior to their decision to enroll at MIT.[21]

Changes in faculty grading behavior would also threaten my assumption of exogeneity. With regards to faculty, I did not have any measures that linked faculty to students in my dataset. Intuitively, the grading-policy change should not have affected faculty behavior, as faculty have

[20] Using a t-test, the pre-policy versus post-policy differences by gender, race/ethnicity, citizenship, and pre-college choice for major were not statistically significant at the .05 level. Student financial need did significantly increase between the two time periods due to increasingly generous financial aid policies. In addition, there were statistically significant differences for the two academic admissions scores (SCORE_OBJ and SCORE_SUBJ). From a practical standpoint, however, these differences were small (average SCORE_OBJ increased by less than .07 points, and average SCORE_OBJ decreased by less than .08 points, both on a 1-5 point scale).

[21] I am not aware of any communications that went to prospective or admitted students that described the change in freshman grading policy.

always assigned a letter grade to student work. What changed was what happened to that letter grade in the registrarial system.

Assuming the policy change was exogenous, below I describe the causal impact of MIT changing from having hidden to externally-shared grades.

Did Changing to Unhidden Grades Affect GPA?

In Table 3, I present descriptive statistics on Freshman GPA in both Fall and Spring for freshmen who entered MIT before and after the freshman grading-policy change. In the fall semester, mean freshman grades differed only slightly pre- and post-policy change. During the pre-policy period, mean GPA ranged between 3.9 and 4.0, while during the post-policy period, mean GPA ranged between 4.0 and 4.1. The fact that there was little difference is expected given that freshmen were graded under identical grading systems in the fall semester, both pre- and post-policy change. Inspecting mean spring semester grades, on the other hand, there was a marked difference. The average spring semester GPA was 3.9 in each of the pre-policy years and 4.2 in each of the post-policy years. This positive and substantial difference in the descriptive data suggests that the freshman grading-policy change impacted freshman academic performance in the spring semester.

In Table 4, I summarize the results of my OLS regression analyses of these same data, with freshman spring semester GPA (SPRINGGPA) treated as the outcome. As indicated in my *Data-Analysis* section, in model specification (1), I fitted a model that contained my key question predictor NOT_HIDDEN and my forcing variable, the year the student entered MIT. In model (2), I added selected student characteristic covariates (dichotomous variables indicating whether the student was female, an underrepresented minority, Asian American, or international) and the financial need of the student. In model (3), I included additional controls representing a student's pre-college ability scores (objective academic score and subjective academic score) and a student's pre-college interest (dichotomous choices for major with "Other" as the reference group). In model (4), I added the two-way interaction between NOT_HIDDEN and a student's pre-college academic performance score, SCORE_OBJ. This model (bolded column in Table 4) is my final model.

In all of the models, as suggested by the simple comparison of means, the estimate on NOT_HIDDEN is statistically significant and positive, suggesting that spring semester freshman year GPAs were higher after grades became externally visible. This effect is robust to the inclusion of various controls. In model specification (4), I examined whether the effect differed by student pre-college academic performance, by including an interaction between NOT_HIDDEN and covariate SCORE_OBJ. The parameter NOT_HIDDEN is now only part of the story.[22] The two-way interaction NOT_HIDDEN * SCORE_OBJ is statistically significant ($p < .05$), meaning that there was a differential effect of eliminating hidden grades on subsequent GPA, by the level of a student's pre-college academic performance. The negative sign of the interaction suggests that those with lower pre-college academic scores experienced a larger gain in spring semester GPA than those with higher scores.

To clarify the impact of this interaction, in Figure 2, I present the fitted relationship between my outcome, freshman spring semester GPA (SPRINGGPA), and the treatment (NOT_HIDDEN) at different levels of a student's prior academic performance (SCORE_OBJ). The solid gray line in Figure 2 shows the predicted values of SPRINGGPA by level of

[22] The presence of this interaction means that the main effect of NOT_HIDDEN represents the causal effect of the grading policy change for students who scored zero on SCORE_OBJ, a condition that does not exist in the sample.

SCORE_OBJ immediately prior to the grading-policy change. The solid black line shows the same relationship immediately after the grading-policy change. As you can see, at each level of SCORE_OBJ, SPRINGGPA is higher for post-policy freshman, compared to pre-policy freshmen. For instance, at the sample average value of SCORE_OBJ (a value of 4.1 on the horizontal axis), I predict that pre-policy freshmen earned a 3.9 GPA, on average, during their spring semester freshman year, controlling for covariates. After the grading-policy change, however, this predicted value increased to 4.2, almost one-third of a point higher. This estimated treatment effect of .31 corresponds, in Figure 2, to the distance along the vertical black dotted line between the fitted plots that I superimposed on the figure at a value of 4.1 for SCORE_OBJ on the horizontal axis.

In addition to this “average” treatment effect, the additional red and green vertical dashed lines that I have superimposed on the figure delineate the respective treatment effects at values of SCORE_OBJ that correspond to the sample mean minus and plus one standard deviation. I have labeled these as “-1 SD” and “+1 SD” values of SCORE_OBJ, respectively. Notice that the treatment effect at the “-1 SD” value of SCORE_OBJ is .35, while at the “+1 SD” value of SCORE_OBJ it is .28. Therefore, the gap between the two regression lines – representing the respective estimated treatment effects – narrows at higher values of SCORE_OBJ. In other words, I have detected a larger impact of the freshman grading-policy change on GPA among students with lower pre-college academic scores than those with higher academic scores.

Did Changing to Unhidden Grades Affect Sophomore Standing?

In Table 5, I show the percent of freshmen who had freshman and sophomore standing in the spring semester by year of entry. As discussed earlier, freshmen are eligible to declare sophomore standing if they complete one quarter of their degree requirements by the end of the fall semester. The primary benefit of being an early sophomore is that a student can declare a departmental major and be assigned a faculty advisor. If a student declares sophomore standing, that student falls under the sophomore grading system with no hidden grades. As shown in Table 5, prior to the grading-policy change, only 1-3% of freshmen in the sample cohorts declared sophomore standing in the spring semester. After the policy change, 9-13% of freshmen declared sophomore standing, a sizable increase.

In Table 6, I summarize the results of my logistic regression analyses with early sophomore standing (SOPHOMORE) treated as the outcome (sophomore standing = 1; freshman standing = 0). I used the same approach to build my regression models for SOPHOMORE as for SPRINGGPA. Again, as suggested by the simple descriptive comparison, the regression analyses confirm that more students declared early sophomore standing after the policy change. This effect is robust to the inclusion of various controls. Model specification (4) contains the two-way interaction between NOT_HIDDEN and SCORE_OBJ. Note, however, that the coefficient on this interaction term is not significant. Therefore, the effect of the grading-policy change on declaration of sophomore status did not vary as a function of student pre-college achievement.

Based on model specification (3), I conclude that the freshman grading-policy change to unhidden grades had a positive effect on the probability of students declaring early sophomore standing, as I found a significant positive coefficient for NOT_HIDDEN ($p < .01$). That is, students had a higher probability of declaring early sophomore standing in their first year spring semester under the new grading policy, compared to the old grading policy. In fact, the fitted odds of a student declaring sophomore standing post-policy was 9.7 times the fitted odds of a student declaring sophomore standing pre-policy.[23] Or, in terms of probabilities, the fitted probability of declaring sophomore standing for post-policy students was seven percent, while the fitted probability of declaring sophomore standing for pre-policy students was less than one percent.

Similar to Figure 2, Figure 3 shows the fitted relationship between my outcome (SOPHOMORE) and the treatment (NOT_HIDDEN) at different levels of a student's prior academic performance (SCORE_OBJ). The solid gray line shows the predicted probabilities of SOPHOMORE immediately prior to the grading-policy change. The solid black line shows the same relationship immediately after the grading-policy change. The vertical black dotted line, between the two solid lines, represents the estimated treatment effect, set at the average value of SCORE_OBJ (a value of 4.1 on the horizontal axis). The difference between where the dotted line intersects the solid black line and the solid gray line is six percentage points (7% probability post-policy minus .8% probability pre-policy).[24]

Did Changing to Unhidden Grades Affect Credit Units Taken?

In Table 7, I show the average total number of credits taken by freshmen during their fall and spring semesters. In the Fall, credits ranged between 54 and 55 both pre-policy and post-policy. Credit units dropped slightly in the Spring across all years, but the numbers were very similar before and after the grading-policy change, suggesting little to no effect of the policy change on credit units taken. This finding is contrary to my hypothesis that students would be

[23] 9.7 is the odds-ratio corresponding to the natural log of the NOT_HIDDEN coefficient in model specification (3) in Table 6, or 2.277.
[24] I only drew a vertical line at the average value of SCORE_OBJ because I do not have an interaction effect between SCORE_OBJ and NOT_HIDDEN. Indeed, if I had plotted the fitted odds ratios, rather than the fitted probabilities, the odds ratio at any value of SCORE_OBJ would be identical.

more likely to reduce their credit load post-policy, compared to pre-policy, since they would face a penalty on the transcript for earning a lower grade.

Recall, however, that post-policy freshmen were more likely to declare early sophomore standing than pre-policy freshmen. In addition, sophomores, unlike freshmen, do not have a cap on the number of credits they can take. Therefore, what I might be seeing is the combined effect of reduced credit hours for some students and increased credit hours for other students, namely those with early sophomore standing[25]. Indeed, when I looked at the distribution of credit units by sophomore standing status, I found that students with early sophomore standing tended to have higher credit unit loads than students with freshman standing. I also found that students who declared early sophomore standing tended to enter MIT with higher levels of SCORE_OBJ.

In Table 8, I build my taxonomy of OLS models in similar fashion to my other research questions, with CREDITS as my outcome variable. Model specification (4) is my final model and includes the two-way interaction between NOT_HIDDEN and a student's pre-college academic performance score, SCORE_OBJ. In this model, the estimate of parameter NOT_HIDDEN is statistically significant ($p < .01$) and negative, meaning that spring semester credits were lower, on average, after grades became externally visible.[26] Since the interaction between NOT_HIDDEN and covariate SCORE_OBJ is statistically significant ($p < .01$), I conclude that there was a differential effect of eliminating hidden grades on subsequent credit units taken, by level of a student's pre-college academic performance. In fact, as I describe below, students who entered MIT with lower levels of SCORE_OBJ tended to slightly decrease

[25] I did not add SOPHOMORE to my regression models to control for early sophomore standing because SOPHOMORE is an outcome variable and is clearly affected by the grading-policy change.
[26] The presence of the interaction NOT_HIDDEN * SCORE_OBJ in the model means that the main effect of NOT_HIDDEN represents the causal effect of the grading policy change for students who scored zero on SCORE_OBJ, a condition that does not exist in the sample.

their credit unit loads, while students who entered MIT with higher levels of SCORE_OBJ tended to slightly increase their credit unit loads.

In Figure 4, I show the effect of this interaction by fitting the relationship between my outcome, freshman spring semester credit units (CREDITS), and the treatment (NOT_HIDDEN) at different levels of a student's prior academic performance (SCORE_OBJ). The solid gray line in Figure 4 shows the predicted values of CREDITS by level of SCORE_OBJ immediately prior to the grading-policy change. The solid black line shows the same relationship immediately after the grading-policy change. At the sample average value of SCORE_OBJ (a value of 4.1 on the horizontal axis), I predict that pre-policy freshmen took 51.5 total credit units, on average, during their spring semester freshman year, controlling for covariates. After the grading-policy change this predicted value decreased very slightly to 51.1 total credit units. This average estimated treatment effect of -.34 credit units corresponds, in Figure 4, to the distance along the vertical black dotted line between the fitted plots that I superimposed on the figure at a value of 4.1 for SCORE_OBJ on the horizontal axis.

In addition to this "average" treatment effect, the additional red and green vertical dashed lines delineate the respective treatment effects at values of SCORE_OBJ that correspond to the sample mean minus and plus one standard deviation, labeled as "-1 SD" and "+1 SD." Notice that the treatment effect at the "-1 SD" value of SCORE_OBJ is -1.15, while at the "+1 SD" value of SCORE_OBJ, it is +.47. This is because the two solid lines in Figure 4 cross. In other words, I have detected a small negative impact of the freshman grading-policy change on subsequent credit units taken among students with lower pre-college academic scores and a small positive effect among students with higher pre-college academic scores.

Did Changing to Unhidden Grades Affect Difficulty of Coursework Taken?

In Table 9, I show the percentage of freshmen who took the standard Physics II course (or equivalent) and the percentage of freshmen who took a more rigorous and mathematically advanced Physics II course. As mentioned earlier, all students at MIT must take or test out of two courses of physics. Most students take Physics II (the second course) in the spring semester. In my sample, 74% of freshmen took a version of Physics II in the Spring. From Table 9, it appears slightly *more* students took the more advanced course after the grading-policy change. Within two years of the cutoff point, 12% of freshmen took the more advanced course pre-policy compared to 13% of students post-policy. This finding is counter-intuitive. When grades are unhidden, I would expect that students would be *less* inclined to enroll in a course that is considered to be more difficult.

I build my taxonomy of regression models in Table 10. Unlike the models for my other outcomes, I did not include controls for a student's *school-level* pre-college major interests. Rather, I included two pre-college *department-level* major controls for interest in Mathematics and Physics, as these two majors correspond more closely with my outcome DIFFICULT. Indeed, when you examine the models that include these controls, the coefficients on PRE_MATH and PRE_PHYS are positive, large, and significant, as one would expect. Model specification (4) is my final model and is bolded in Table 10. This model includes the two-way interaction NOT_HIDDEN * SCORE_OBJ.

In model (4), the estimate of parameter NOT_HIDDEN is statistically significant and positive, as is the interaction between NOT_HIDDEN and SCORE_OBJ.[27] The significant (p <

[27] The presence of this interaction means that the main effect of NOT_HIDDEN represents the causal effect of the grading policy change for students who scored zero on SCORE_OBJ, a condition that does not exist in the sample.

.01) two-way interaction NOT_HIDDEN * SCORE_OBJ means that there was a differential effect of eliminating hidden grades on the probability of taking the more rigorous version of Physics II by level of a student's pre-college academic performance. In fact, students who entered MIT with lower levels of SCORE_OBJ had a lower probability of taking the more difficult Physics II course, while students who entered MIT with higher levels of SCORE_OBJ had a higher probability of taking the same course.

To clarify the impact of this interaction, in Figure 5, I present the fitted relationship between my outcome, taking a more mathematically advanced Physics II course in the spring semester (DIFFICULT), and the treatment (NOT_HIDDEN) at different levels of a student's prior academic performance (SCORE_OBJ). The solid gray line in Figure 5 shows the predicted values of DIFFICULT by level of SCORE_OBJ immediately prior to the grading-policy change. The solid black line shows the same relationship immediately after the grading-policy change. At the sample average value of SCORE_OBJ (a value of 4.1 on the horizontal axis), I predict that pre-policy freshmen had a 8.5% probability of taking the more rigorous Physics II course, on average, during their spring semester freshman year, controlling for covariates. After the grading-policy change, this predicted value decreased slightly to 7.4%. This estimated treatment effect of minus one percentage point corresponds, in Figure 5, to the distance along the vertical black dotted line between the fitted plots that I superimposed on the figure at a value of 4.1 for SCORE_OBJ on the horizontal axis. Put another way, the fitted odds of a student taking the more rigorous Physics II course post-policy was .86 times the fitted odds of a student taking this same course pre-policy.

In addition to this "average" treatment effect, the additional red and green vertical dashed lines that I have superimposed on the figure delineate the respective treatment effects at values of

SCORE_OBJ that correspond to the sample mean minus and plus one standard deviation. I have labeled these as “-1 SD” and “+1 SD” values of SCORE_OBJ, respectively. Notice that the treatment effect at the “-1 SD” value of SCORE_OBJ is a negative two percentage points, while at the “+1 SD” value of SCORE_OBJ, it is a positive three percentage points. The change in sign is a result of the two solid lines in Figure 5 crossing.

It is worth noting that the more advanced Physics II course did not substantively change in terms of content or pedagogy in the years of my sample. However, the standard Physics II course did. Specifically, the teaching format for the standard course changed from lecture/recitation to a more collaborative, technology-driven format in the same year as the grading-policy change. In addition, in the early years of my sample, some students took an experimental version of the standard Physics II course using a hands-on/lab-based format. A potential threat to validity is that some students who would have taken the standard Physics II course, but did not like interactive learning pedagogies, may have chosen to take the more advanced Physics II course instead. Therefore, these estimates should be interpreted with this fact in mind.

Conclusions and Limitations

This study examined the effects of changing from hidden to externally-visible grades during the second semester of freshman year at MIT. Using the policy change as a natural experiment, I presented causal estimates that addressed the selection issues that have plagued the previous literature on the effects of grading policies on subsequent student performance and behavior. I found that the grading-policy change had several important effects on students.

First, the elimination of hidden grades in the second semester resulted in higher freshman GPA. Specifically, post-policy freshmen earned grades that were nearly one-third of a point higher than pre-policy freshmen. Assuming grades signify, at least partly, the extent to which students master course material, this is a positive result – particularly in light of the fact that students often take foundational courses during their first year, and these courses presumably help students in their transition to their majors and to life beyond college. Interestingly, one of the original motivations for having hidden grades was to provide freshmen a suitable environment, free of undue pressure and anxiety, to acquire foundational knowledge. If grades are an appropriate measure of a student's acquisition of foundational knowledge, it appears that, at least in the second semester, freshmen are better served by having their passing grades appear on the transcript. Another potentially favorable finding is that the grading-policy change had the largest positive impact on those students who entered MIT with lower test scores and grades. In other words, the grading-policy change seemed to encourage better performance among those less well-prepared, compared to students who started college with higher levels of academic scores.

The mechanism by which grades increased as a result of the freshman grading-policy change is not entirely clear. Intuitively, freshmen may have worked harder in their classes when they knew that the grades that they received would be recorded and shared with external parties. In other words, hidden grades may have previously acted as a disincentive for students to take their coursework seriously. Nevertheless, there may be other explanations. For example, freshmen may have reduced the number of credit units taken spring semester to allow for more time to devote to their coursework. Indeed, I found that freshmen tended, on average, to take a slightly lighter course load under the post-policy grading system, when spring semester grades were unhidden, compared to the pre-policy grading system.[28] This difference, however, is quite small. Given that freshmen typically take a 50-plus credit load in a semester, a 0.3 credit unit difference, while statistically significant, does not hold much practical significance. I also found an interaction effect between the grading-policy change and a student's level of pre-college academic performance. Specifically, students who entered MIT with lower academic credentials tended, on average, to drop one credit unit in the spring, while students with higher academic credentials tended to add an additional one-half credit unit. Therefore, the reduction in credit hours seemed to be a mechanism exercised primarily by students with lower pre-college performance.

The change in freshman grading policy also had a positive effect on the probability of freshmen declaring early sophomore standing in their spring semester. In fact, the odds of post-policy freshmen declaring sophomore standing was almost ten times the odds for pre-policy freshmen. Given that hidden grades in the spring semester were eliminated post-policy, students who were eligible to become early sophomores in the Spring had little incentive to remain freshmen and great incentive to begin work on their majors as sophomores. Based on the data available in the study, there were no negative consequences to more freshmen declaring early sophomore standing as a result of the grading-policy change. In fact, early sophomores, both pre- and post-policy, tended to enter MIT with stellar academic credentials and, while at MIT, seemed to exercise their intellectual interests to their fullest.

Finally, I found that freshmen were slightly less likely, on average, to take a more rigorous, mathematically advanced second course in Physics in the post-policy period than the

[28] I do not know if the positive impact of the grading policy change on spring semester GPA was due specifically to freshmen taking fewer credit units. I cannot answer this question based on the analytical approach I used in this study. In my model for SPRINGGPA, I did *not* include a control for CREDITS, because CREDITS is an outcome variable.

pre-policy period, controlling for covariates. Nevertheless, the average probability of taking the more advanced version of Physics II differed by just one percentage point between pre-policy freshmen and post-policy freshmen. Therefore, while students may have chosen their coursework, in small part, based on whether or not the grade would appear on the official academic transcript, the form of grading policy was likely not the primary motivation for this behavior. More interestingly, I found that the grading-policy shift impacted students differently, depending on how well they performed academically in high school and on various pre-college standardized tests. Specifically, freshmen with lower pre-college academic scores were less likely to take the more rigorous course post-policy, compared to pre-policy, while students who entered MIT with the highest level of academic scores tended to do the opposite.

Most earlier studies of this topic have attempted to eliminate bias in estimates of the effect of a grading-policy change on student outcomes, which exists due to the endogeneity of the grade choice, by introducing controls for observable characteristics, such as pre-college student measures. However, such studies cannot support unbiased causal conclusions, as they do not account for unobserved influences on the grade choice, such as a student's own expectations of performance. While I argue that my estimates of the impact of the grading-policy change on four outcomes are causal, given the exogenous assignment of freshmen to a particular grading system, I recognize that there are several potential threats to the validity of my findings, which I have noted throughout this paper.

One additional limitation of my study was the presence of another condition on campus that shifted at the same time as the grading-policy change. Specifically, prior to fall 2002, one-third of freshmen at MIT opted to live off-campus. Beginning in fall 2002, with the construction of a new undergraduate residence hall, all freshmen were required to live on-campus. The simultaneous introduction of a new freshman housing policy and a new freshman grading policy makes it difficult to isolate the degree to which the grading-policy change influenced my outcomes versus the change in housing policy. In an attempt to better understand the effects of the change in grading policy alone, I replicated all of my analyses in a restricted sample of freshmen who lived on-campus before and after the policy change. My findings remained the same using this limited sample. While using a limited sample is a simple strategy for checking the sensitivity of my findings, I understand that these results are potentially biased, as I have not accounted for why pre-policy students chose to live on- or off-campus. Still, my results conform to what one would predict based on theory, lending credibility to my argument that the change in grading policy had important effects on student behaviors.

References

Arcidiacono, Peter (2003). Ability sorting and the returns to college major. Duke University, Department of Economics, Working Paper, 02-26.

Bedard, Kelly (2001). Human capital versus signaling models: university access and high school dropouts. *The Journal of Political Economy*, 109(4), 749-775.

Camara, Wayne J. and Gary Echternacht (2000). The SAT I and high school grades: Utility in predicting success in college. College Board Report No. RN-10. New York: College Board.

Cherry, Todd L. and Larry V. Ellis (2005). Does rank-order grading improve student performance? Evidence from a classroom experiment. *International Review of Economics Education*, 4(1), 9-19.

Giometti, Thomas (1976). One experience with the pass-fail grading system. *Hispania*, 59(2), 302-307.

Guskey, Thomas R. and Jane M. Bailey (2001). *Developing grading and reporting systems for student learning*. Thousand Oaks, CA: Corwin Press, Inc.

Harris, Gregory A. (2009). The impact of hidden grades on student academic performance: A case study on a single university campus. Harvard University Graduate School of Education.

Koenker, Roger and Kevin F. Hallock (2001). Quantile regression. *Journal of Economic Perspectives*, 15(4), 143-156.

Light, Richard J., Judith D. Singer, and John D. Willett (1990). *By design: Planning research on higher education*. Cambridge, MA: Harvard University Press.

Michaels, James W. (1977). Classroom reward structures and academic performance. *American Educational Research Association*, 47(1), 87-98.

MIT (2005). *2005 alumni survey results*. Retrieved from (URL hidden) on February 22, 2008.

MIT (2000, September). Freshman pass/no record grading and advanced placement policy. Report of a special subcommittee on the Committee on the Undergraduate Program. Retrieved from (URL hidden) on December 19, 2007.

MIT (2008, January). Review of pass/no record grading and the sophomore exploratory subject. Report of the subcommittee of the Committee on the Undergraduate Program. Retrieved from (URL hidden) on February 22, 2008.

Montmarquette, Claude, Kathy Cannings, and Sophie Mahseredjian (2002). How do young people choose college majors? *Economics of Education Review*, 21, 543-556.

Olsen, Turee (1975). ERIC/RCS report: grading alternatives. *The English Journal*, 64(3), 106-108.

Rask, Kevin N. and Elizabeth M. Bailey (2002). Are faculty role models? Evidence from major choice in an undergraduate institution. *Journal of Economic Education*, 33(2), 99-124.

Rask, Kevin N. and Jill Tiefenthaler (2004). Too few women or too many men? The gender imbalance in undergraduate economics. http://ssrn.com/abstract=595222.

Rohe, Daniel E. et al. (2006). The benefits of pass-fail grading on stress, mood, and group cohesion in medical students. *Mayo Clinic Proceedings*, 81(11), 1443-1448.

Sabot, Richard and John Wakeman-Linn (1991). Grade inflation and course choice. *The Journal of Economic Perspectives*, 5(1), 159-170.

Sgan, Mathew R. (1970). Letter grade achievement in pass-fail courses. *The Journal of Higher Education*, 41(8), 638-644.

Stallings, William M. and H. Richard Smock (1971). The pass-fail grading option at a state university: a five semester evaluation. *Journal of Educational Measurement*, 8(3), 153-160.

Wittich, Barbara Von (1972). The impact of the pass-fail system upon achievement of college students. *The Journal of Higher Education*, 43(6), 499-508.

Figure 1. Type of freshman grading policy by entry year, 1998-2005.

	Pre-Policy Period				Post-Policy Period			
Fall	**Pass/No-Record (HIDDEN)**				**Pass/No-Record (HIDDEN)**			
Spring	**Pass/No-Record (HIDDEN)**				**A/B/C/No-Record (UNHIDDEN)**			
Entering Fall:	1998	1999	2000	2001	2002	2003	2004	2005
Observations:	1,029	1,038	999	1,016	963	1,010	1,060	983

Figure 2. Fitted values of SPRINGGPA by freshman grading policy (pre vs. post) and SCORE_OBJ using data from Model (4) of the fitted taxonomy of OLS regression models in Table 4. Covariates are set at their mean values.

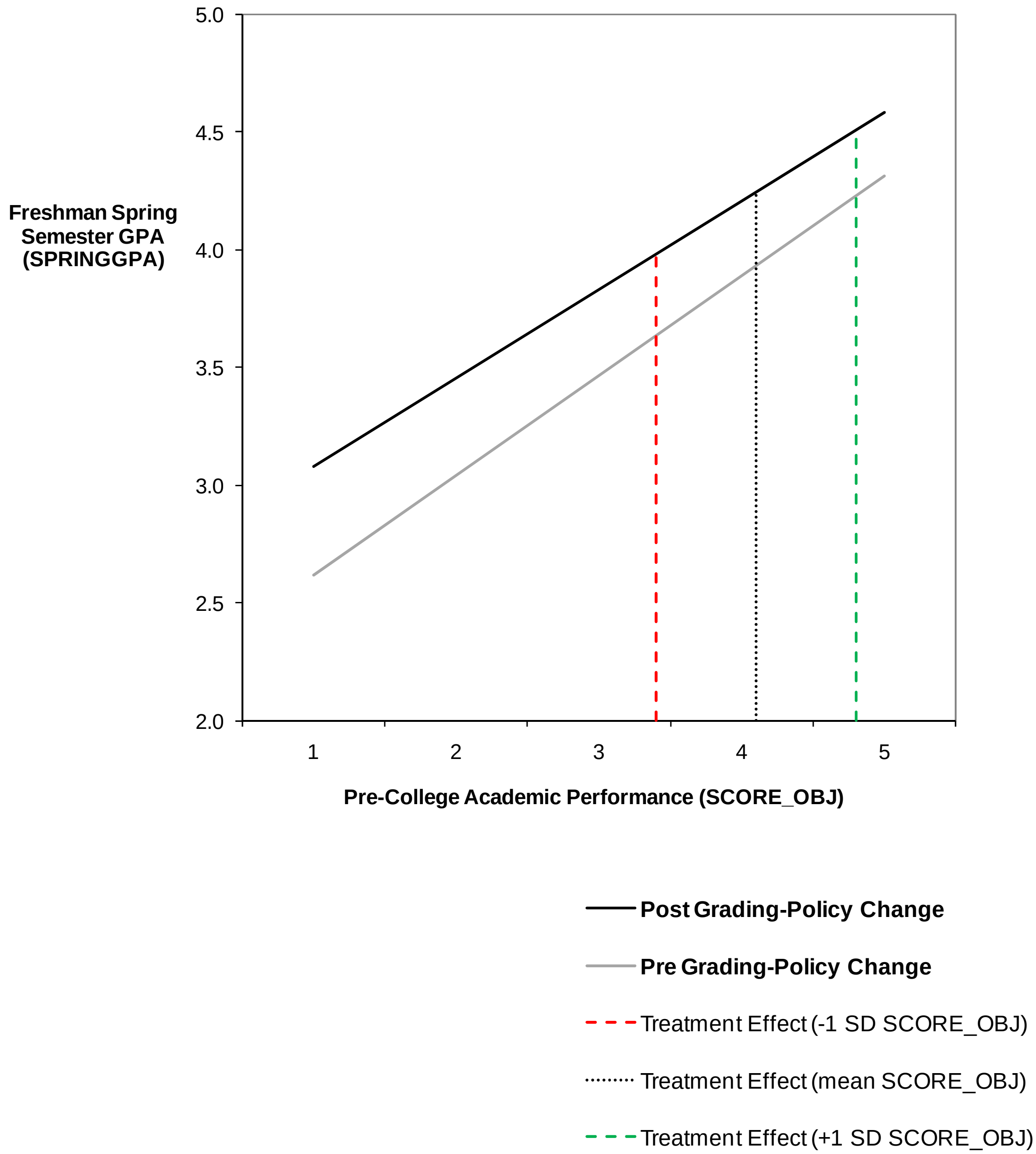

Figure 3. Fitted probability of SOPHOMORE by freshman grading policy (pre vs. post) and SCORE_OBJ using data from Model (3) of the fitted taxonomy of logistic regression models in Table 6. Covariates are set at their mean values.

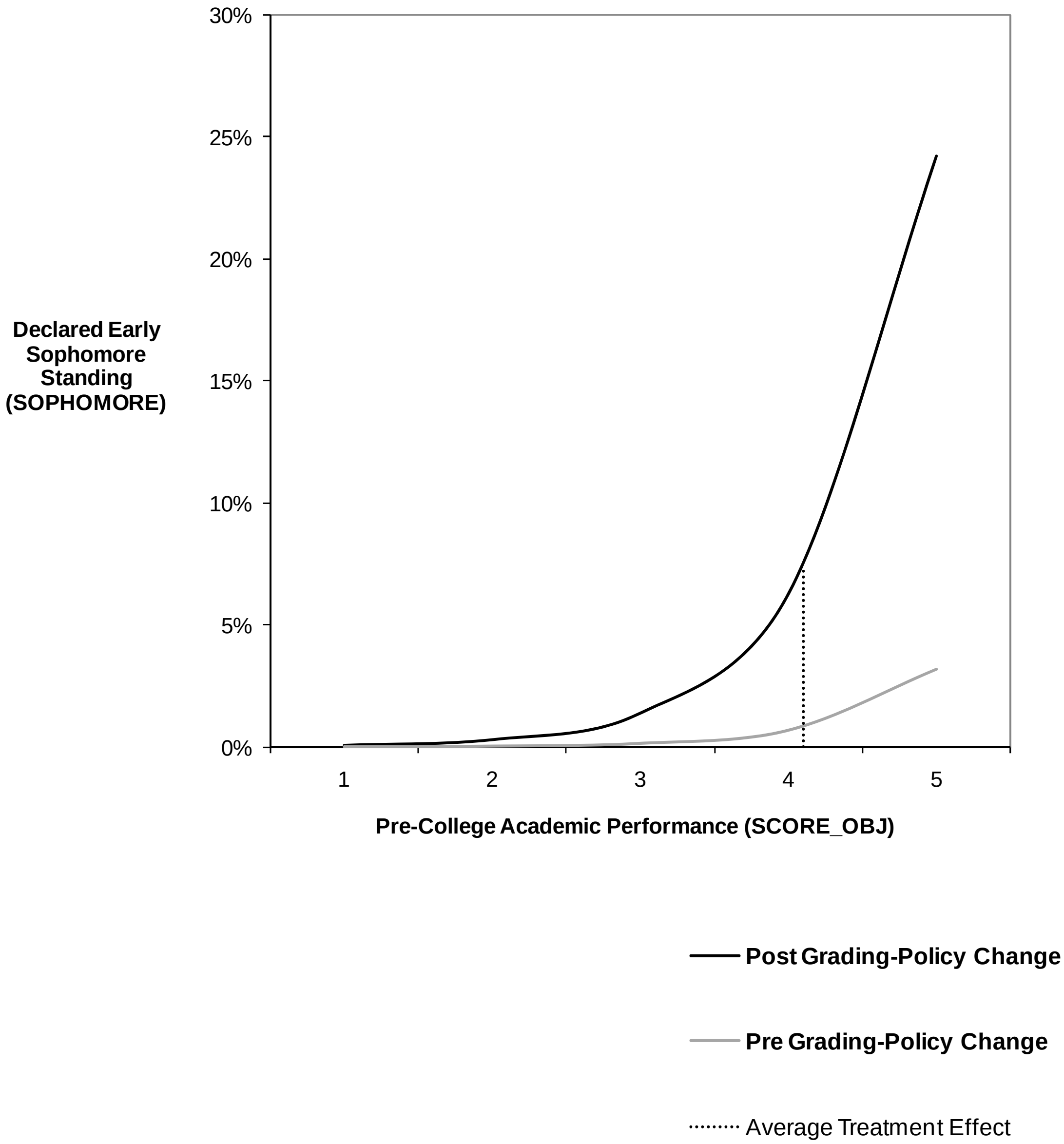

Figure 4. Fitted values of CREDITS by freshman grading policy (pre vs. post) and SCORE_OBJ using data from Model (4) of the fitted taxonomy of OLS regression models in Table 8. Covariates are set at their mean values.

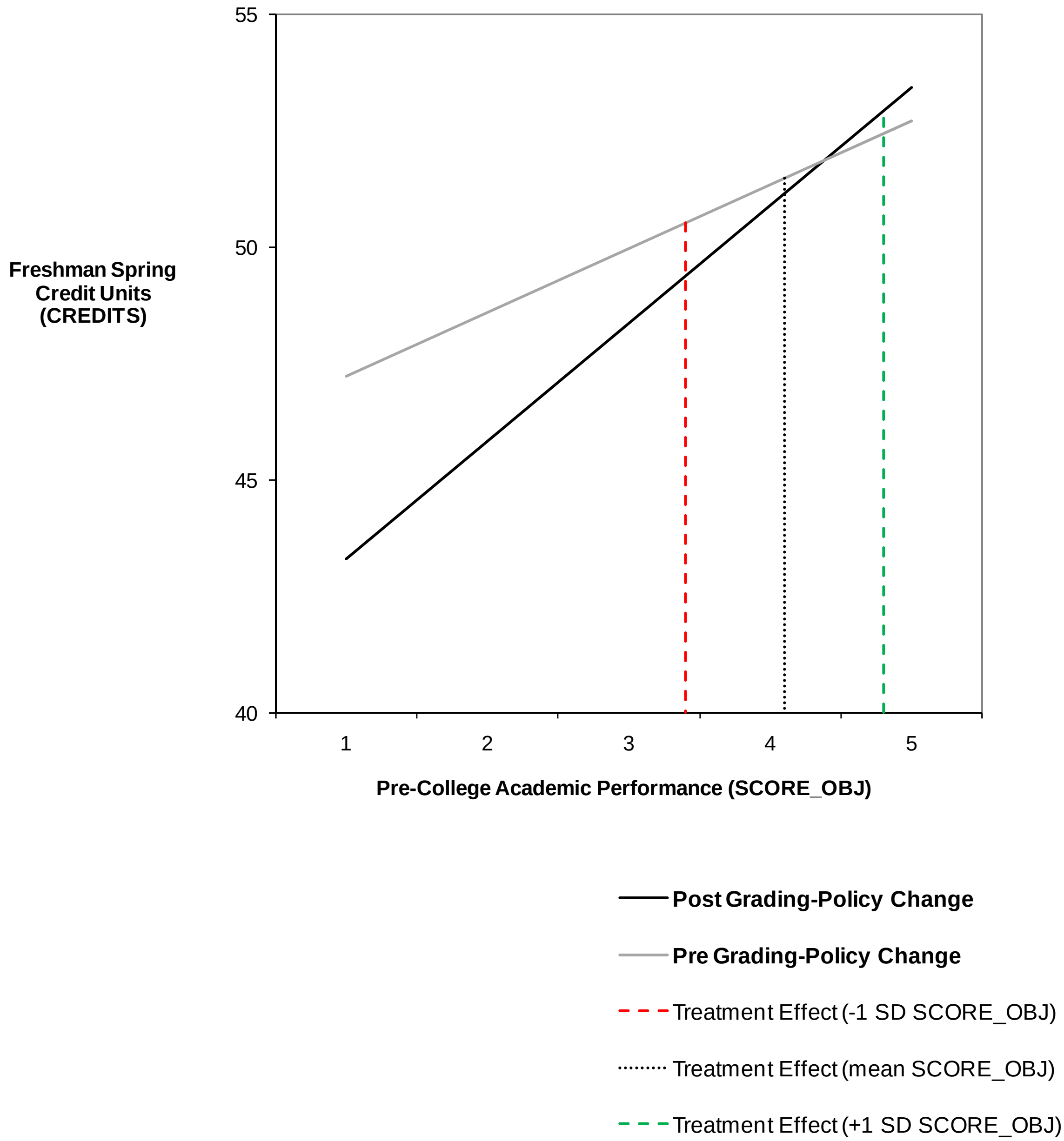

Figure 5. Fitted probability of DIFFICULT by freshman grading policy (pre vs. post) and SCORE_OBJ using data from Model (4) of the fitted taxonomy of logistic regression models in Table 10. Covariates are set at their mean values.

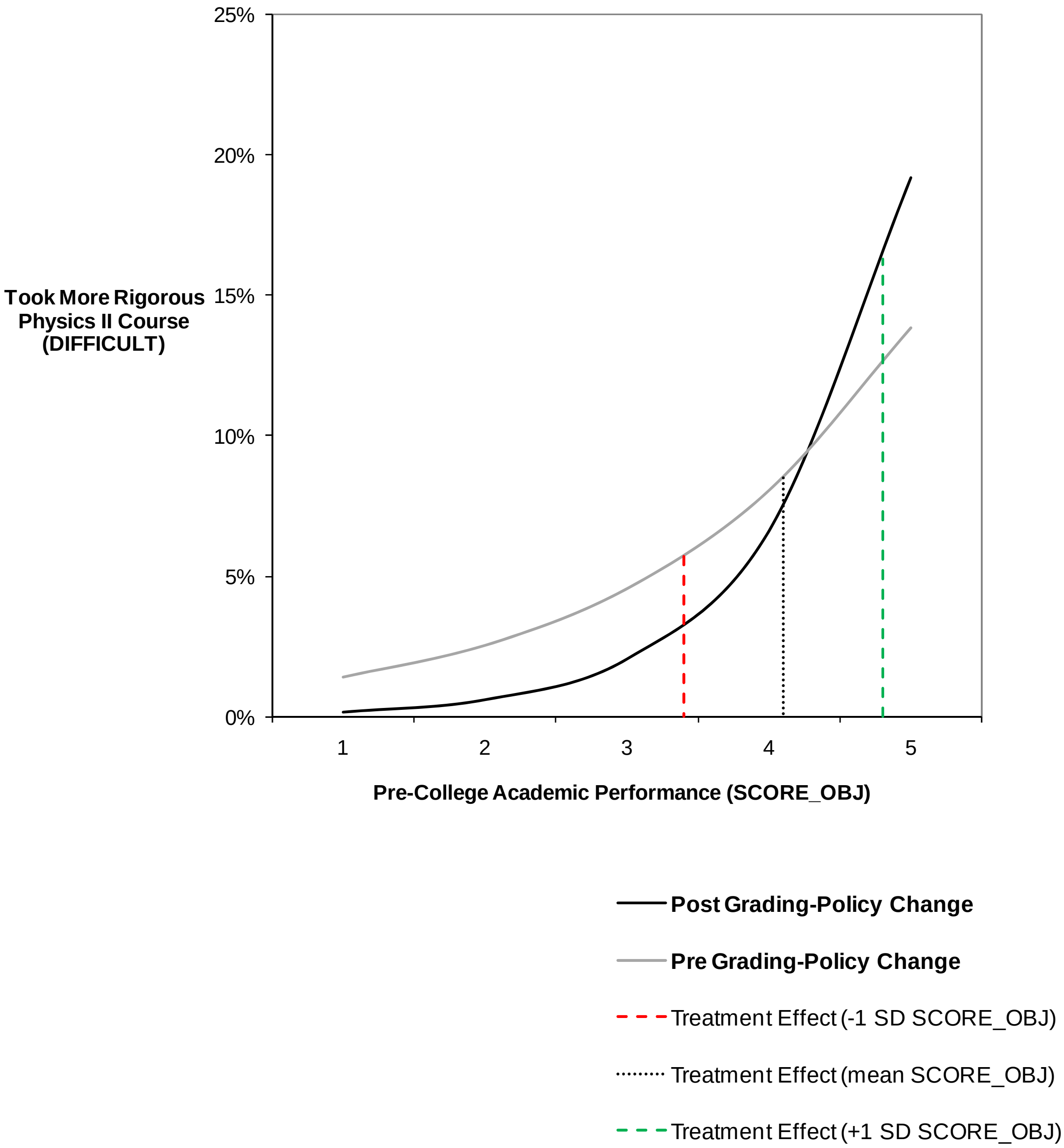

Table 1: Names, Definitions, and Coding of Variables in the Research

Category	Variable	Definition	Value Range
Outcomes	SPRINGGPA	Student's individual grade point average freshman year second semester.	0-5.
	CREDITS	Student's total credit units freshman year second semester.	12-102.
	SOPHOMORE	Student declared sophomore standing freshman year second semester.	1=sophomore standing; 0=freshman standing.
	DIFFICULT	Student took more mathematically advanced Physics II course freshman year second semester. This variable is only defined for students who took Physics II spring semester.	1=took advanced Physics II; 0=took standard Physics II.
Question predictor	NOT_HIDDEN	Student did not receive hidden grades second semester freshman year	1=yes (post-policy); 0=no (pre-policy).
Student demographics	YEAR	Year student entered MIT.	Integers, centered on year in which grading policy changed (-3; -2; -1; 0; 1; 2; 3; 4)
	FEMALE	Female dummy variable.	1=female; 0=male.
	URM	Underrepresented minority dummy variable (American Indian or Alaskan Native, Black or African American, Hispanic or Latino).	1=underrepresented minority; 0=otherwise.
	ASIAN	Asian American dummy variable.	1=asian american; 0=otherwise
	INTERNTL	International dummy variable.	1=international; 0=domestic.
	NEED	Financial need = cost of attendance minus expected family contribution, as determined by institutional methodology.	0-$52,473.
Student pre-college choice for major	PRE_ARCH	Pre-college choice for major is in School of Architecture (reference school = Other).	1=school of architecture; 0=otherwise.
	PRE_ENG	Pre-college choice for major is in School of Engineering (reference school = Other).	1=school of engineering; 0=otherwise.
	PRE_HUM	Pre-college choice for major is in School of Humanities, Arts, and Social Sciences (reference school = Other).	1=school of humanities, arts, and social sciences; 0=otherwise.
	PRE_MGT	Pre-college choice for major is in School of Management (reference school = Other).	1=school of management; 0=otherwise.
	PRE_SCI	Pre-college choice for major is in School of Science (reference school = Other).	1=school of science; 0=otherwise.

Category	Variable	Definition	Value Range
	PRE_PHYS	Pre-college choice for major is Physics. *Note: This variable is used as a control only for outcome DIFFICULT.*	1=physics; 0=otherwise.
	PRE_MATH	Pre-college choice for major is Math. *Note: This variable is used as a control only for outcome DIFFICULT.*	1=math; 0=otherwise.
Student pre-college performance	SCORE_OBJ	Admissions objective academic score combining math SAT, verbal SAT, science SAT, high school grades, and high school rank.	1-5.
	SCORE_SUBJ	Admissions subjective academic score regarding student motivation, sense of purpose, level of high school courses, and commitment.	1-5.

Table 2: Summary statistics of variables used in analyses, N = 8,098.

Name	Description	Pre-Policy Change (Entered Fall 1998-2001)			Post-Policy Change (Entered Fall 2002-2005)		
		Mean	**Std. Dev.**	**Range**	**Mean**	**Std. Dev.**	**Range**
NOT_HIDDEN	Student did not receive hidden grades second semester freshman year (=1).	0.00	0.00	0 - 0	1.00	0.00	1 - 1
SPRINGGPA	Student's individual grade point average freshman year second semester.	3.90	0.79	0 - 5	4.21	0.72	0 - 5
CREDITS	Student's total credit units freshman year second semester.	51.37	5.78	12 - 99	51.06	7.18	24 - 102
SOPHOMORE	Student declared sophomore standing freshman year second semester (=1).	0.02	0.14	0 - 1	0.11	0.31	0 - 1
DIFFICULT	Student took more mathematically advanced Physics II course freshman year second semester (=1). This variable is only defined for students who took Physics II spring semester.	0.11	0.31	0 - 1	0.14	0.34	0 - 1
YEAR	Year student entered MIT, centered on year in which grading policy changed.	N/A	N/A	-3 - 0	N/A	N/A	1 - 4
FEMALE	Female (=1).	0.42	0.49	0 - 1	0.44	0.50	0 - 1
URM	Underrepresented minority (=1).	0.20	0.40	0 - 1	0.18	0.39	0 - 1
ASIAN	Asian American (=1).	0.28	0.45	0 - 1	0.27	0.44	0 - 1
INTERNTL	International (=1).	0.07	0.26	0 - 1	0.07	0.26	0 - 1
NEED	Financial need = cost of attendance minus expected family contribution, as determined by institutional methodology (2005 dollars)	14,227	15,156	0 - 61,131	17,488	16,809	0 - 50,350
PRE_ARCH	Pre-college choice: Architecture major (=1).	0.02	0.12	0 - 1	0.01	0.12	0 - 1
PRE_ENG	Pre-college choice: Engineering major (=1).	0.56	0.50	0 - 1	0.57	0.50	0 - 1

Name	Description	Pre-Policy Change (Entered Fall 1998-2001) Mean	Std. Dev.	Range	Post-Policy Change (Entered Fall 2002-2005) Mean	Std. Dev.	Range
PRE_HUM	Pre-college choice: Humanities, Arts, or Social Sciences major (=1).	0.03	0.18	0 - 1	0.03	0.17	0 - 1
PRE_MGT	Pre-college choice: Management major (=1).	0.02	0.14	0 - 1	0.02	0.14	0 - 1
PRE_SCI	Pre-college choice: Science major (=1).	0.34	0.47	0 - 1	0.35	0.48	0 - 1
PRE_PHYS	Pre-college choice: Physics major (=1).	0.07	0.25	0 - 1	0.08	0.28	0 - 1
PRE_MATH	Pre-college choice: Math major (=1).	0.08	0.27	0 - 1	0.08	0.27	0 - 1
SCORE_OBJ	Admissions objective academic score combining math SAT, verbal SAT, science SAT, high school grades, and high school rank.	4.08	0.71	1 - 5	4.15	0.65	1.8 - 5
SCORE_SUBJ	Admissions subjective academic score regarding student motivation, sense of purpose, level of high school courses, and commitment.	3.55	0.64	2 - 5	3.47	0.61	2 - 5
Observations			4,082			4,016	

Notes: The sample is freshmen who entered MIT Fall 1998-2005 with information from administrative files. Observations for DIFFICULT are 2,488 (pre-policy) and 2,521 (post-policy). This is because the analysis for this outcome is restricted to freshmen who took Physics II spring semester.

Table 3: Sample average freshman GPA for fall and spring semester, by year, pre- and post-policy change.

	Fall				Spring			
Entering Fall:	Mean	Std. Dev.	Range	N	Mean	Std. Dev.	Range	N
1998	3.9	0.8	0 - 5	1,029	**3.9**	0.8	0 - 5	1,029
1999	3.9	0.8	0 - 5	1,038	**3.9**	0.8	0 - 5	1,038
2000	4.0	0.8	0 - 5	999	**3.9**	0.8	0.7 - 5	999
2001	4.0	0.7	0 - 5	1,016	**3.9**	0.8	0 - 5	1,016
2002	4.0	0.7	0 - 5	963	**4.2**	0.7	0 - 5	963
2003	4.0	0.7	1 - 5	1,010	**4.2**	0.7	0 - 5	1,010
2004	4.0	0.8	0 - 5	1,060	**4.2**	0.7	0 - 5	1,060
2005	4.1	0.7	0 - 5	983	**4.2**	0.7	0 - 5	983

Note: The dotted line represents the freshman grading-policy change. Pre-policy years, when second semester grades were hidden, are above the dotted line. Post-policy years, when second semester grades were not hidden, are below the dotted line.

Table 4: Parameter estimates, approximate p-values, and goodness-of-fit statistics for a taxonomy of fitted ordinary least-squares regression models describing the relationship between a freshman grading-policy change and freshman spring semester GPA **(SPRINGGPA)**.

	Baseline Model	Adding Student Characteristics	Adding Pre-College Ability/ Interest	**Adding Interaction with SCORE_OBJ**
Specification	(1)	(2)	(3)	**(4)**
Post grading-policy change (NOT_HIDDEN)	.256** (.035)	.261** (.034)	.310** (.032)	**.506** (.100)**
Year student entered MIT, centered (YEAR)	.015* (.008)	.016* (.007)	-.002 (.007)	**-.003 (.007)**
Female (FEMALE)		-.074** (.017)	-.012 (.016)	**-.012 (.016)**
Underrepresented minority (URM)		-.408** (.023)	-.115** (.024)	**-.114** (.024)**
Asian American (ASIAN)		-.004 (.020)	-.084** (.019)	**-.083** (.019)**
International (INTERNTL)		.222** (.033)	.185** (.032)	**.184** (.032)**
Financial need (NEED)		-.000** (.000)	-.000** (.000)	**-.000** (.000)**
Objective academic score (SCORE_OBJ)			.402** (.013)	**.423** (.017)**
Subjective academic score (SCORE_SUBJ)			.057** (.013)	**.058** (.013)**
Pre-college choice for Architecture (PRE_ARCH)			.157* (.078)	**.159* (.078)**
Pre-college choice for Engineering (PRE_ENG)			.074 (.049)	**.075 (.049)**
Pre-college choice for Humanities (PRE_HUM)			-.028 (.064)	**-.024 (.064)**
Pre-college choice for Management (PRE_MGT)			-.081 (.071)	**-.080 (.071)**
Pre-college choice for Science (PRE_SCI)			.061 (.049)	**.063 (.049)**
NOT_HIDDEN * SCORE_OBJ Interaction				**-.047* (.023)**
Intercept	3.919** (.016)	4.063** (.020)	2.059** (.083)	**1.968** (.094)**
R^2 statistic	0.043	0.106	0.210	**0.210**
SS_{Model}	207.433	511.115	1009.637	**1011.652**
SS_{Error}	4604.864	4301.181	3802.659	**3800.644**
Observations	8,098	8,098	8,098	**8,098**

Standard errors in parentheses. ˜p<.10 * p<.05 ** p<.01

Table 5: Sample proportion of students with freshman standing and sophomore standing spring semester, by year, pre- and post-policy.

	Freshman Standing		Sophomore Standing	
Entering Fall:	N	%	N	%
1998	1,009	98%	20	**2%**
1999	1,012	97%	26	**3%**
2000	981	98%	18	**2%**
2001	1,001	99%	15	**1%**
2002	842	87%	121	**13%**
2003	918	91%	92	**9%**
2004	946	89%	114	**11%**
2005	888	90%	95	**10%**

Notes: The dotted line represents the freshman grading-policy change. Pre-policy years, when second semester grades were hidden, are above the dotted line. Post-policy years, when second semester grades were not hidden, are below the dotted line.

Before spring semester, freshmen are notified by the Registrar if they are eligible for early sophomore standing. Eligibility is based mostly on percent completion of general institute requirements. The primary benefit of being an early sophomore is that a student can declare a departmental major and be assigned a faculty advisor. If a student declares sophomore standing, that student falls under the sophomore grading system.

Table 6: Parameter estimates, approximate p-values, and goodness-of-fit statistics for a taxonomy of fitted logistic regression models describing the probability of a freshman declaring early sophomore standing **(SOPHOMORE)**.

	Intercept Only	Baseline Model	Adding Student Characteristics	**Adding Pre-College Ability/ Interest**	Adding Interaction with SCORE_OBJ
Specification	(0)	(1)	(2)	**(3)**	(4)
Post grading-policy change (NOT_HIDDEN)		2.106** (.212)	2.110** (.213)	**2.277** (.223)**	1.697 (1.408)
Year student entered MIT, centered (YEAR)		-.080~ (.042)	-.057 (.043)	**-.079~ (.046)**	-.079~ (.046)
Female (FEMALE)			-.776** (.104)	**-.351** (.112)**	-.351** (.112)
Underrepresented minority (URM)			-.828** (.190)	**.221 (.204)**	.222 (.204)
Asian American (ASIAN)			.861** (.106)	**.537** (.113)**	.537** (.113)
International (INTERNTL)			.271 (.186)	**-.359~ (.203)**	-.359~ (.203)
Financial need (NEED)			.000** (.000)	**.000* (.000)**	.000* (.000)
Objective academic score (SCORE_OBJ)				**1.565** (.126)**	1.463** (.275)
Subjective academic score (SCORE_SUBJ)				**.938** (.082)**	.938** (.082)
Pre-college choice for Architecture (PRE_ARCH)				**-.751 (1.153)**	-.751 (1.153)
Pre-college choice for Engineering (PRE_ENG)				**.825 (.529)**	.827 (.530)
Pre-college choice for Humanities (PRE_HUM)				**.452 (.629)**	.452 (.630)
Pre-college choice for Management (PRE_MGT)				**.658 (.651)**	.661 (.652)

	Intercept Only	Baseline Model	Adding Student Characteristics	**Adding Pre-College Ability/ Interest**	Adding Interaction with SCORE_OBJ
Specification	(0)	(1)	(2)	**(3)**	(4)
Pre-college choice for Science (PRE_SCI)				**1.074*** **(.530)**	1.076* (.530)
NOT_HIDDEN * SCORE_OBJ Interaction					.126 (.302)
Intercept	-2.719** (.046)	-4.050** (.132)	-3.844** (.150)	**-15.355**** **(.843)**	-14.884** (1.402)
X^2 statistic	3758.677	3475.775	3301.906	**2860.561**	2860.390
ΔX^2 from model (0)		282.902	456.771	**898.116**	898.287
Pseudo-R^2 statistic		.075	.122	**.239**	0.239
Observations	8,098	8,098	8,098	**8,098**	8,098

Standard errors in parentheses. ~p<.10 * p<.05 ** p<.01

Table 7: Sample average freshman credit units for fall and spring semester, by year, pre- and post-policy change.

	Fall				Spring			
Entering Fall:	Mean	SD	Range	N	Mean	SD	Range	N
1998	54.9	6.0	24 - 72	1,029	**51.3**	5.8	24 - 99	1,029
1999	53.9	6.0	30 - 91	1,038	**51.4**	5.7	24 - 81	1,038
2000	55.2	6.2	30 - 72	999	**51.4**	6.0	12 - 81	999
2001	55.1	6.4	25 - 72	1,016	**51.4**	5.6	24 - 72	1,016
2002	54.9	6.5	36 - 72	963	**51.4**	6.7	24 - 87	963
2003	54.5	6.4	24 - 72	1,010	**50.7**	7.3	24 - 102	1,010
2004	54.4	6.3	24 - 70	1,060	**50.7**	7.4	24 - 96	1,060
2005	54.1	6.5	30 - 73	983	**51.5**	7.3	24 - 102	983

Notes: The dotted line represents the freshman grading-policy change. Pre-policy years, when second semester grades were hidden, are above the dotted line. Post-policy years, when second semester grades were not hidden, are below the dotted line.

Credit units for a course are defined as the total number of hours spent each week in class/laboratory plus the estimated time that the average student spends each week in outside preparation for that course.

Table 8: Parameter estimates, approximate p-values, and goodness-of-fit statistics for a taxonomy of fitted ordinary least-squares regression models describing the relationship between a freshman grading-policy change and freshman spring semester credit units **(CREDITS)**.

	Baseline Model	Adding Student Characteristics	Adding Pre-College Ability/ Interest	**Adding Interaction with SCORE_OBJ**
Specification	(1)	(2)	(3)	**(4)**
Post grading-policy change (NOT_HIDDEN)	-.426 (.299)	-.405 (.296)	-.252 (.289)	**-5.094** (.912)**
Year student entered MIT, centered (YEAR)	.029 (.065)	.036 (.064)	-.018 (.063)	**-.002 (.063)**
Female (FEMALE)		-1.001** (.148)	-.482** (.149)	**-.472** (.148)**
Underrepresented minority (URM)		-1.857** (.206)	-.425~ (.218)	**-.457* (.218)**
Asian American (ASIAN)		.701** (.174)	.258 (.172)	**.244 (.172)**
International (INTERNTL)		1.118** (.289)	.438 (.293)	**.456 (.292)**
Financial need (NEED)		-.000** (.000)	-.000* (.000)	**-.000~ (.000)**
Objective academic score (SCORE_OBJ)			1.891** (.121)	**1.370** (.152)**
Subjective academic score (SCORE_SUBJ)			1.039** (.122)	**1.020** (.122)**
Pre-college choice for Architecture (PRE_ARCH)			.870 (.717)	**.832 (.716)**
Pre-college choice for Engineering (PRE_ENG)			1.100* (.447)	**1.073* (.446)**
Pre-college choice for Humanities (PRE_HUM)			-.338 (.589)	**-.425 (.588)**
Pre-college choice for Management (PRE_MGT)			.380 (.654)	**.361 (.652)**
Pre-college choice for Science (PRE_SCI)			.514 (.452)	**.463 (.451)**
NOT_HIDDEN * SCORE_OBJ Interaction				**1.160** (.207)**
Intercept	51.415** (.142)	52.127** (.180)	39.474** (.764)	**41.711** (.861)**
R^2 statistic	0.001	0.026	0.073	**0.077**
SS_{Model}	201.151	9088.150	25100.942	**26330.316**
SS_{Error}	343478.030	334591.030	318578.239	**317348.865**
Observations	8,098	8,098	8,098	**8,098**

Standard errors in parentheses. ~p<.10 * p<.05 ** p<.01

Table 9: Of freshmen who took Physics II spring semester, sample proportion who took the standard version and the mathematically advanced version, by year, pre- and post-policy change.

	Physics II (standard)		**Physics II (mathematically advanced)**	
Entering Fall:	N	%	N	%
1998	719	92%	63	**8%**
1999	707	90%	81	**10%**
2000	638	88%	89	**12%**
2001	677	88%	91	**12%**
2002	628	87%	95	**13%**
2003	652	87%	97	**13%**
2004	669	87%	96	**13%**
2005	604	84%	112	**16%**

Notes: The dotted line represents the freshman grading-policy change. Pre-policy years, when second semester grades were hidden, are above the dotted line. Post-policy years, when second semester grades were not hidden, are below the dotted line.
The analysis for this research question is limited to students who took a version of Physics II spring semester.

Table 10: Parameter estimates, approximate p-values, and goodness-of-fit statistics for a taxonomy of fitted logistic regression models describing the probability of a freshman taking a more mathematically advanced Physics II course spring semester **(DIFFICULT)**.

	Intercept Only	Baseline Model	Adding Student Characteristics	Adding Pre-College Ability/ Interest	**Adding Interaction with SCORE_OBJ**
Specification	(0)	(1)	(2)	(3)	**(4)**
Post grading-policy change (NOT_HIDDEN)		-.109 (.164)	-.127 (.167)	.005 (.176)	**-2.607**** **(.706)**
Year student entered MIT, centered (YEAR)		.097** (.036)	.099** (.036)	.075~ (.039)	**.078*** **(.039)**
Female (FEMALE)			-.794** (.088)	-.653** (.093)	**-.658**** **(.093)**
Underrepresented minority (URM)			-.926** (.129)	-.246~ (.141)	**-.251~** **(.141)**
Asian American (ASIAN)			-.523** (.105)	-.638** (.111)	**-.644**** **(.112)**
International (INTERNTL)			.465** (.146)	.357* (.162)	**.346*** **(.163)**
Financial need (NEED)			.000 (.000)	.000 (.000)	**.000~** **(.000)**
Objective academic score (SCORE_OBJ)				.894** (.085)	**.608**** **(.109)**
Subjective academic score (SCORE_SUBJ)				.483** (.072)	**.476**** **(.072)**
Pre-college choice for Physics (PRE_PHYS)				1.735** (.121)	**1.737**** **(.121)**
Pre-college choice for Math (PRE_MATH)				1.215** (.130)	**1.211**** **(.131)**
NOT_HIDDEN * SCORE_OBJ Interaction					**.599**** **(.157)**
Intercept	-1.990** (.040)	-1.992** (.078)	-1.499** (.096)	-7.498** (.450)	**-6.238**** **(.537)**
X^2 statistic	4423.636	4403.572	4213.281	3747.353	**3732.699**
ΔX^2 from model (0)		20.064	210.355	676.283	**690.937**
Pseudo-R^2 statistic		0.005	0.048	0.153	**0.156**
Observations	6,018	6,018	6,018	6,018	**6,018**

Standard errors in parentheses. ~p<.10 * p<.05 ** p<.01

Note: The analysis for this research question is limited to students who took a version of Physics II spring semester.