



Steven J. Green
School of International
& Public Affairs

Department of Economics

I Promise to Work Hard: The Impact of Non-Binding Commitment Pledge on Academic Performance

Nicholas A. Wright

Puneet Arora

Jesse Wright

Department of Economics

Florida International University

Working Paper 2411

September 2024

11200 SW 8th Street, Miami, Florida 33199

<https://economics.fiu.edu/>

I Promise to Work Hard: The Impact of a Non-Binding Commitment Pledge on Academic Performance*

Nicholas A. Wright[†]
Florida International University

Puneet Arora
Management Development Institute, Gurgaon, India

Jesse Wright
Florida Gulf Coast University

Abstract

Students often start a course with high expectations and an ambitious plan of action. Some instructors use goal-inducing non-binding commitment pledges to nudge students to follow through on their intended course of action. Using a field experiment, we asked treated students to set a goal grade, identify the actions they will take to achieve it, and sign a commitment pledge to work towards this grade. We find that while treated students pledged a greater time commitment and targeted a higher grade, their overall test scores decrease by 0.23 standard deviations and they were 15 percentage points less likely to pass the course.

Keywords: Commitment Pledge; Goal-setting; Academic Performance; Overconfidence; Grade Expectations

JEL Codes: A2, C93, D8, I2

*We would like to thank Augustine Denteh, Ishita Tripathi, and seminar participants at Florida Gulf Coast University for their insightful comments on a previous version of this paper. All remaining errors are our own. The AEA registry RCT number for this study is AEARCTR-0008173 and the Florida Gulf Coast University IRB Protocol number is 2021-66. Declarations of interest: none.

[†]Corresponding author: Department of Economics, Florida International University, Miami, FL. E-mail: nicwri@fiu.edu.

1 Introduction

Students often enter a course with high hopes of success and an ambitious plan of action. However, all too often, these expectations are not met, and academic performance and effort can fall far short. Low academic performance can have a significantly negative effect on engagement and psychological well-being (Wrosch et al., 2003), self-regulated learning (Stone, 2000), subsequent performance (Rask, 2010), university dropout rates (Stinebrickner and Stinebrickner, 2014), and long-term goal attainment (Ordóñez et al., 2009; Nowell and Alston, 2007). To combat these effects, policymakers have implemented costly interventions, such as merit-based financial aid and scholarships (Andrews et al., 2020; Page et al., 2019; Clotfelter et al., 2018), individualized student coaching (Bettinger and Baker, 2014), performance standard policies (Wright, 2020; Lindo et al., 2010), and tutoring (Hardt et al., 2023; Fryer Jr, 2017; Cook et al., 2014). This paper examines the effect of a more economical alternative: having students state a target grade, identify the effort to achieve it, and then sign a non-binding commitment pledge to work towards it. Research on the causal impact of such academic pledges on student performance is surprisingly limited considering their widespread use in education.

In this paper, we investigate the combined impact of goal setting and a non-binding commitment pledge on student performance in an introductory economics course. We conduct a field experiment in which students are randomly assigned to either a treatment or control group. The control group students disclose their expected grade from the course. The treatment group, however, not only disclose their expected grade but also set goals by revealing their preferred grade, identify the effort needed to achieve it, and sign a non-binding commitment pledge to the instructor to signify their dedication to working towards their preferred grade. We then analyze the effects of this intervention on student performance on two midterm exams, assignment submission and scores, final exam, and their final course grade. While our focus is on academic performance, it is worth noting that such soft commitment devices, as defined in Bryan et al. (2010), are prevalent across academic institutions, where

students make pledges related to various aspects, ranging from class attendance to timely assignment submissions, and from plagiarism to upholding integrity and ethics.

Research in the field of education has consistently shown the positive effects of goal-setting on student performance. By setting reference points, goal-setting can motivate hyperbolic discounters or loss-averse students to put in higher effort to meet the set objectives (Koch and Nafziger, 2016, 2011). These studies have demonstrated that goal-setting can have a positive impact on performance in various settings. For instance, setting goals based on accomplishment in an entire course have been shown to improve student performance (van Lent and Souverijn, 2020; Koch and Nafziger, 2020; Schippers et al., 2015; Morisano et al., 2010), and setting goals based on an individual task in the course have a similar effect (Clark et al., 2020). However, more recent studies have shown that combining goal-setting with text message reminders has no effect and that setting goals too high may even harm student performance (Oreopoulos et al., 2022; Dobronyi et al., 2019; van Lent and Souverijn, 2020).¹ Overall, research on goal-setting in the field of education has produced mixed results, and further exploration is still necessary.

Similarly, non-binding commitment pledges have been studied in several contexts including loan repayment, tax declaration, sex abstinence, volunteerism, and charity contribution or in-kind donations (Bearman and Brückner, 2001; Kettle et al., 2017; Bhanot, 2017; Andreoni and Serra-Garcia, 2021a; Capra et al., 2022; Cotterill et al., 2013; Meyer and Tripodi, 2021). Most studies have found that such commitment pledges have a positive or null impact on these outcomes. Positive impacts are attributed to the pledger’s preference to keep their promise (Vanberg, 2008) or to their guilt-averse attitude (Charness and Dufwenberg, 2006), while null effects are attributed to time-inconsistent preferences where the pledger later reneges on their non-binding promise (Laibson, 1997; Andreoni and Serra-Garcia, 2021b).² In

¹See Ordóñez et al. (2009) for a discussion of several scenarios that may cause goal-setting to have adverse consequences in different decision making contexts.

²Reneging is a low-cost option for non-binding pledges. Studies show that pledgers renege on their commitment to donate money to a charity, blood, or books to a library, work as a volunteer, and repay their loans on time (Bhanot, 2017; Capra et al., 2022; Cotterill et al., 2013; Fosgaard and Soetevent, 2018; Meyer and Tripodi, 2021).

some cases, pledges can even have an unintended negative effect due to the frustration caused by the pledger’s inability to honor their pledge (Sutan et al., 2018; Oppenheimer and Olivola, 2011).³ Although non-binding commitment pledges are widely used in education, there have been relatively few studies examining their effect on student outcomes. While Himmler et al. (2019) investigates the effects of a soft commitment device in education, wherein students sign a nonbinding agreement and commit to staying on track for graduation, their design does not include goal-setting. Their findings indicate that those offered the soft commitment device were more likely to sign up for, participate in, and pass exams.

Theoretically, both goal-setting and commitment pledge are likely to help students overcome their self-control problems, and may result in a stronger combined positive impact. However, the education context is unique compared to other pledge settings, as students have a lower degree of control over the fulfillment of their pledge. When a student promises to achieve a particular course grade, for instance, their ability, prior achievement, effort, peer network, family inputs, and coping strategies all contribute to the outcome, many of which the student has limited or no control over (Hanushek, 2010). As such, when combined with goal setting, a commitment pledge may lead to overconfident students locking themselves into an ambitious goal they cannot achieve. While a student may renege on their non-binding promise, the potential disappointment associated with not meeting their targeted outcome may have adverse effects. Thus, it is important to understand the true causal impact of goal-inducing non-binding commitment pledges on student performance.

Our results show that while the treated students pledged a greater time commitment and a higher targeted grade, the intervention decreased their overall test score by 0.23 standard deviations and reduced the likelihood of passing the course by 15 percentage points. We further confirm these results using the randomization inference test, which shows that our point estimates are very robust and unlikely to be obtained by chance. These findings are similar to the results of Koch and Nafziger (2020) and van Lent and Souverijn (2020), which

³See Bryan et al. (2010) for a review of studies based on commitment devices.

shows that the effect of goal-setting can be reversed when it is externally imposed by the experimenter (even when the external interference is minimal) or the goal-setter is induced to target a more ambitious goal. In our context, similar to [van Lent and Souverijn \(2020\)](#), students are also induced to set a higher goal grade than they expected, which may partially explain the negative effects we found.

A study that closely aligns with our design is [Goldhaber et al. \(2020\)](#), which evaluates a non-binding commitment to excel in high school, be a responsible citizen, avoid felony convictions, and apply for financial aid. Similar to our findings, their results indicate effects that are either insignificant or negative. However, distinctions between their study and ours are evident: our investigation involves a non-incentivized pledge, while theirs is incentivized with potential financial rewards upon pledge fulfillment. Our study involves self-selected individual goal while theirs is a state-declared goal applicable to every student. Moreover, our study is an experimentally examined impact, contrasting with their intent-to-treat analysis. Other studies explore the impact of binding commitment pledges within goal-setting environments. For example, [Patterson \(2018\)](#) examined the influence on performance in an online course when students committed to limiting their daily internet time using software to a self-selected threshold, beyond which all sites would be blocked. Their findings show a positive effect for that intervention. On the contrary, [Ariely and Wertenbroch \(2002\)](#) allowed students to self-impose deadlines for turning in their exam papers and found that they committed to early deadlines (analogous to a higher preferred grade in our setting) but performed worse than students on whom those deadlines were externally imposed and evenly spaced (analogous to a lower likelihood of passing in our setting). In our study, however, we utilized a non-binding commitment device, which differs in nature from the binding commitments used in [Ariely and Wertenbroch \(2002\)](#) and [Patterson \(2018\)](#).

Our work broadly contributes to three streams of literature. First, our research is part of the growing literature in the behavioral economics of education, which uses low-cost interventions to address the deficits in students non-cognitive skills. This literature has found

that such interventions can effectively improve student performance and motivation.⁴ Second, despite the prominent combined use of (soft) non-binding pledges within a goal-setting context, our paper is the first to examine their causal impact on college students' performance. Lastly, we contribute to a growing strand of the economics literature which finds that because of time-inconsistent preferences, commitment pledges may yield unintended effects. More generally, the findings from our study may offer important insights about the use of non-binding commitment pledges in other contexts where the pledger has lesser control over the fulfillment of their promise.

2 The Experimental and Empirical Design

Our field experiment was conducted in an introductory economics course at a public university in the Southwestern United States. This university is a medium-sized institution with a total enrolment of about 15,000 students. To be eligible for admission, students must have an SAT total of at least 1070 and an ACT composite score of 21 and above. Among the Freshman Class of 2021, the mean SAT and ACT scores were 1141 and 23.4, respectively. On average, the ability of the students in our sample is representative of the pool of admitted students, with a mean SAT score of 1144.8 and a composite ACT score of 23.3.⁵ These students are also quite motivated and ambitious, with 68% reporting that they were ranked in the top 25% of their high school class. In addition, the typical student planned to devote about 5.3 hours studying and completing assignments in the course each week.

⁴See [Damgaard and Nielsen \(2018\)](#) for a review of these works. Studies have used low or zero cost behavioral interventions to address sub-optimal performance problems. Such interventions include presenting points on a test as points lost (vs. points gained) ([McEvoy et al., 2016](#)), providing relative performance information feedback in a relative grading system ([Kajitani et al., 2020](#)) or an absolute grading system ([Tran and Zeckhauser, 2012](#)), increasing the frequency of examinations ([De Paola and Scoppa, 2011](#); [Tuckman, 1998](#)), using text messages to provide simplified academic information, encouragement, or reminders ([Oreopoulos et al., 2022](#); [Castleman and Meyer, 2020](#)), or providing information about past performance in the course ([Wright and Arora, 2022](#)). Research has shown that these interventions can effectively improve student performance and motivation. However, there are also other low-cost behavioral interventions in education that have been found to be ineffective ([Krawczyk et al., 2011](#); [Balart et al., 2022](#)).

⁵These scores are just slightly above the mean SAT scores of 1051 to 1068 for the cohort of high school students who took the exam between 2017 and 2020 ([National Center for Education Statistics, 2021](#)).

About 40% of students drop, fail, or withdraw (DFW) from this introductory course each semester. The high DFW rate is concerning because this is a general education course that is taken by students across the university and a gateway course for several business majors. One of the main factors contributing to this high DFW rate is that students likely overestimate their expected performance in the course. This is because: (i) information about the past distribution of grades is not publicly available, and open sources (e.g., Rate My Professor) provide inaccurate information about this distribution, (ii) the students who enroll may not be interested in economics or they may lack the necessary preparation and subject-specific ability, and (iii) newer college students (Freshmen and Sophomores) are more than two-thirds of enrolled students.⁶ Since students hold inaccurate expectations of their academic performance, this is an interesting context to study the impact of commitment pledges as it may induce students to exert more effort over the semester.

In this paper, we examine the impact of an experimental intervention that randomly selects a group of students to sign an academic pledge to work towards their preferred grade and identify the time they promise to invest to achieve the targeted grade. The experiment was executed in 4 Principles of Macroeconomics sections that were offered by two male instructors in Fall 2021.⁷ The students who agreed to participate in the study were required to complete an online survey during the second week of classes - after students were no longer able to add or drop courses without penalty. As such, irrespective of their treatment status, all participating students were required to provide information about their socioeconomic attributes, study habits, past academic outcomes, risk attitudes, and predicted grades for

⁶Independent of their cognitive ability, students are likely to underperform in general education courses. [Thompson et al. \(2015\)](#) argue that because these courses are mandatory, they do not usually align with students' interests or future career aspirations, and as such, enrolled students may have lower subject-specific ability relative to a course they self-selected. In addition, they found that most students would not enroll in general education classes if given a choice. This is one potential explanation for the high DFW rate.

⁷3 of the 4 classes were in-person, and 1 was delivered online. While there may be small variations, the course policy is highly standardized across instructors. Across all courses, the grading components can be summarized as follows: midterm exams (55-60%), final exam (30-35%), homework assignments (5%), and attendance/participation (5%). Because of the sample size limitations, we are unable to assess if the mode of delivery affects the estimates.

the course. A copy of the survey is available in Online Appendix A.⁸

After completing the general baseline questions, students were then randomly assigned to the treatment and control groups. For students assigned to the control group, the survey ended immediately and they continued the semester as they otherwise would. On the other hand, the treated students were required to make a promise to their professor and themselves that they will take several productive steps to achieve a target grade, one that they then select. The targeted grade can be higher, lower, or the same as the grade they predicted earlier in the survey. We then ask treated students to choose the percent of classes they promise to attend, the weekly number of hours they promise to devote to quizzes or homework assignments, and the weekly number of hours they promise to allocate to reading for the course, in their quest to attain the pledged grade. As such, all treated students are simultaneously making a pledge to target a certain level of effort (input) and a particular outcome (output) in the course.⁹

2.1 Data and Randomization

The data we utilize in this study comes from two main sources: the baseline survey and students' official academic records. Table 1 shows that the average age of study participants was 20.05, 63% were males, 21% were first-generation college-goers, and 69% were either Freshmen or Sophomores. In addition, about 43% of students had a household income below sixty thousand dollars, 66% of students were employed, and the typical student worked about

⁸While the spillover of information between treated and untreated students is possible, there are two main reasons why this is unlikely in our study. First, since the survey was administered online, and there were only minor differences in the survey seen by students in the treatment and control groups, there is less opportunity for students to observe and share the differences in the survey they took. Second, while the sample size is significantly reduced, we find similar results when we analyze the data by in-person and online classes. This is important because there are fewer connections and less communication between students in online classes.

⁹Regardless of their decision to participate in the study, all students were awarded an extra credit of less than 1%, based on the proportion of students who completed the survey in each classroom. As such, opting out of the survey reduces the extra credit that is awarded to all students in the class. Of the 218 enrolled students, 82.6% or 180 students agreed to participate and completed the baseline survey, but only 164 of these students completed the course (i.e. drop-out rate of 8.9%). We find no evidence of selection in the decision to withdraw from the course.

13.5 hours. At the start of the course, students average GPA was 3.35 and about 60% of the students in the sample passed the math pre-test. Lastly, the average test scores on midterm 1, midterm 2 and the final exam were 73.75, 63.79, and 65.63, respectively.

The table also shows that there are no statistical differences in the average pre-intervention characteristics of students in the treated and control groups. As such, there is a high degree of confidence in our random assignment of students to treatment and any observed differences in students' course outcomes must be due to the intervention. While less than 20% of students typically obtain an A- or A in this course, Table 10 in Online Appendix A shows that 41% of the treated students expected to receive this grade, and about 53% of treated students were pledging to work towards it. In addition, the table shows that treated students revised their time commitment upward after making the pledge, and students were more likely to pledge a higher grade than what they predicted they would earn in the course.

2.2 Empirical Design

We utilize the following empirical model to estimate the causal impact of our intervention:

$$Y_i = \alpha_0 + \alpha_1 P_i + X_i \beta + \gamma_p + \gamma_s + \epsilon_i,$$

where Y_i is the outcome of interest for student i , including students' test scores on the midterm and final exams, their overall course score, and a pass/fail indicator variable. P_i is an indicator variable that takes the value 1 if the student is randomly assigned to the treatment group and X_i is a vector of control variables including various pre-treatment attributes such as students' age, gender, current GPA, mathematics test score, typical study hours, labor market engagement, and first-generation status. In addition, all models include instructor (γ_p) and section (γ_s) fixed effects. The coefficient of interest, α_1 , shows the causal effect of the treatment on students' academic outcomes in the course.¹⁰ In addition to the model

¹⁰The instructors taught multiple sections. Removing the covariates, instructor, and section fixed effects do not change the magnitude of the point estimates, though they become slightly noisier. In addition, including measures of risk preference did not change the main conclusion.

above, we also utilize a similarly specified ordered Probit model to examine the impact of the intervention on students predicted grade and actual final grade at the end of the course.

With a sample of 180 observations, we have found very precise estimates of the treatment effect of the intervention.¹¹ To assess the implications of this limitation, we also provide p-values from a two-sided randomization inference test of no treatment effect. This procedure randomly reassigns students in each classroom to a placebo control/treatment group 1000 times - estimating a placebo treatment effect each time. The randomized inference p-value is then calculated as the share of placebo treatment effects that is greater than the baseline point estimates in absolute terms (Fujiwara and Wantchekon, 2013).

3 Empirical Results

In Table 2, we show the impact of the intervention on students' pre-treatment predicted grades, the pledged grade of treated students relative to the benchmark predicted grade, and students' actual grades at the end of the course. As expected, the estimates in column 1 show that students in the treatment and control groups had a similar prediction of their final grade before the intervention. Column 2 then shows that the fraction of students who promised to target an A- or A was 11 percentage points (pp) higher than the baseline share of students who predicted that they would earn these grades. However, while treated students were very aspirational when making their pledges, column 3 shows that they were 7pp less likely to achieve a grade of A- or A, 6pp less likely to earn a B-, B, or B+, and 11pp more likely to fall below a grade of D+.

Table 3 shows the impact of the intervention on students' test scores and their overall course performance. The outcomes of interest that are presented in columns 1-4 are average

¹¹We initially planned to repeat the experiment across several semesters, targeting at least 500 observations and 10 classrooms to be able to detect a small effect size with .8 power. However, given the large and precise negative effects we observed after the first iteration of the intervention, we decided to end the experiment early due to the ethical concerns this raises and the potential harm the treatment imposes on treated participants. With an equally allocated sample size of 180 and .8 power, we can detect an effect size of 0.38 and above. While we find slightly smaller effects, we are not concerned about power because all our main results are statistically significant and precisely estimated.

mid-term score, final exam score, overall course score, and a pass/fail indicator variable. We also present the heteroskedasticity-robust standard error in round brackets and the randomization inference p-value in square brackets below each estimate. The results indicate that the intervention reduced treated students' average mid-term score by 0.14 standard deviations (SD), their final exam score by 0.28SD, and their overall course score by 0.23SD. In addition, the students who made the performance pledge were less likely to pass the course by 15 percentage points (pp) or 24% of the pass rate in the control group.

To assess how students are adjusting their behavior during the semester, we examine the impact of the intervention on mid-terms 1 and 2 separately, assignment scores, the likelihood of missing assignments, and the likelihood of withdrawing from the course. These results are presented in Table 4. The estimates suggest that the intervention had no impact on test scores for the first mid-term exam, but it reduced test scores by 0.27SD on the second midterm exam and average homework scores by 0.28SD. However, positive insignificant effects were found for the likelihood of missing a homework assignment or withdrawing from the course. This evidence suggests that the performance of treated students started to deteriorate after they failed to meet their pledged grade on their first assessment.

3.1 Prediction Errors and Students Subsequent Performance

The baseline results revealed that there is no difference in the performance of treated and untreated students in the first midterm assessment. However, roughly 65% of all students scored more than 5% below the threshold for their predicted course grade, and half of the students scored more than 10% below that threshold. We hypothesize that treated students reacted more negatively to the feedback they received on their first assessment because the commitment pledge increased the salience of failing to meet their targeted grade.

3.1.1 Descriptive Trends

To explore this further, Figure 1-3 show how students' prediction error at the first mid-term exam impacted their subsequent performance on later assessments.¹² On each graph, the vertical axis shows the degree to which a student's performance has improved on a later assessment compared to their performance on the first midterm exam, and the horizontal axis shows the degree to which each student's first midterm exam score deviated from their predicted course grade (prediction error).

Figure 1 shows that in the second mid-term exam, both treated and untreated students generally received scores that were lower than their first mid-term test score. However, treated students experienced a larger decline in their subsequent performance relative to untreated students. This result is consistent for the students who failed to meet, met, or exceeded their expectation in the first mid-term exam, and the negative gains on exam 2 fade out linearly as the prediction error increases – i.e. for the more overconfident or highly ambitious students. Figure 2 presents a similar outcome, with both treated and untreated students experiencing negative net gains on the final exam relative to their first mid-term exam. However, treated students still generally performed worse than untreated students. Similar to Figure 1, this negative effect also tends toward zero as students' distance from their predicted grade on the first mid-term exam increases.

Lastly, in Figure 3, we examine how the prediction error impacted students' final score in the course. This final assessment is a comprehensive measure of student's overall effort in the class because it includes their homework, class participation, and all their exam scores. Similar to the previous results, we find that treated students under-performed their untreated counterparts across the entire prediction error distribution. While it appears that students with high prediction error on midterm 1 made greater improvements on subsequent

¹²In the baseline survey, all students were asked to reveal their predicted letter grade for the course. To calculate the prediction error, we converted the predicted letter grades into the minimum percentage score that each student needed to realize their prediction. We then minus the student's actual score on the first mid-term exam.

assessments, the observed improvement is greater among the untreated students. This is clear as the prediction error threshold that corresponds to positive gains on subsequent assessments is at least twice as high for treated students, relative to the threshold for untreated students.

3.1.2 Mechanism and Mediation Analysis

Since the trends in Figures 1-3 are descriptive, we next estimate how students prediction error on their first exam interacted with the treatment to impact their subsequent performance. These results are presented in Table 5. To create this table, we first separate the prediction error distribution into quartiles and then conduct subgroup analysis across the resulting 4 groups. The estimates for each quartile is presented in columns 1-4, respectively. For the first quartile, the range of the prediction error variable shows that students mostly met or exceeded their predictions, with actual midterm scores being as much as 7.7 percentage points higher and as worse as as 2.74 percentage points lower than students' prediction. Similarly, the scores in the second, third, and fourth quartiles were 2.75-12.04, 12.05-20.94, and 20.95-53.55 percentage points below students' prediction, respectively. The outcome variables in rows 1-3 are the net change in test scores between the first midterm exam and students performance on midterm 2, the final exam, and their overall course grade, respectively. The results suggest show that the negative effects of the treatment on students subsequent performance is being driven by the treated students who slightly or moderately under-performed their predicted level of performance on the first mid-term exam. In contrast, the treated students on the extreme who over-performed, met or grossly under-performed their predicted performance continued to maintain the same level of performance as the control group on subsequent course assessments. For these students, the intervention had no meaningful impact on their academic outcome in the course.

Lastly, we utilize mediation analysis to isolate how the direct effect changes when known mediators are included in the model.¹³ For this analysis we focus on students' overall course

¹³In this research, our primary focus is on examining the impact of a pledge intervention (X) on student performance (Y). We have posited that students may experience disappointment and subsequently alter their

grade, but the analysis could easily be repeated for other outcomes. These results are presented in Table 6. While the treatment had no impact on students' performance on the first mid-term exam, the results show that controlling for prediction error (i.e. predicted score-exam 1 score) significantly reduce the performance gap in the course. When we also control for various measures of student effort, such as the percent of missed homework assignments and homework scores, the treatment gap in overall performance closes by about 70%, from 0.23 SD to an insignificant 0.07SD. As such, the negative effect is partly explained by the fact that treated students were induced to exert less effort in the course. Lastly, controlling for test scores on the second mid-term exam fully erases the impact of the treatment on students' overall course performance.

Similarly, in Table 2O of the online appendix, we examine how various course assessment mediate the effect of treatment on the overall course score. These results also show that the first mid-term exam only mediated 4% of the treatment effect. However, including any subsequent assessment (homework assignments, midterm exam 2, or final exam) was sufficient to drive the treatment effect to zero. As such, the treatment affected students' final course outcome through their performance on the assessments completed after their first exam.

Together, these results support our hypothesis that one plausible explanation for the unintended negative impact of commitment pledges on students subsequent performance is that the commitment pledge paired with goal setting may induce a large discouragement effect after an initial setback. That is, after receiving low scores in their first assessment, students may foresee that they are unable to achieve the ambitious grade they promised. This may

behavior in response to the feedback they receive on their first exam. If this hypothesis holds true, it follows that the intervention should have a discernible effect on students' behavior and subsequent performance on future exams. As such, we aim to examine how the intervention's impact on these subsequent outcomes (M) explains their overall performance in the course. This evaluation can be executed through a three-step process. First, we estimate the overall impact by regressing the treatment on the overall course grade, taking into account only pre-treatment covariates. Next, we estimate the treatment's effect on the potential mediating variables, such as midterm exam results, final exam scores, homework scores, and the likelihood of missing exams (see Tables 3 & 4). Lastly, we revisit the initial step, this time incorporating the mediating variables. If the treatment's effect diminishes when M is included in the model, it suggests that this variable serves as a mediator in explaining how the treatment impacts students' overall course grades.

cause discouragement and lead to undesirable changes in their subsequent behavior. The results show that both groups first mid-term scores were far below their expectations and that both groups experienced a negative gain in their subsequent performance. However, the setback was more salient and the discouragement effect stronger for the treated students who made the commitment pledge to pursue a particular grade. This explains the negative treatment effect we observe for the intervention group on subsequent exams. Similarly, several studies have found that the students who grossly overestimated their academic performance in the first year of college get into an academic trap where they are subsequently more likely to procrastinate, cram for exams, lower their expectations, and study less (Beattie et al., 2018, 2019). Our results are also consistent with van Lent and Souverijn (2020) who found a negative effect of goal-setting when mentors advised students to set a higher goal than what the student initially selected. Given the novelty of our findings, we implore future studies to continue building on our results and to further explore this discouragement effect mechanism in education and other contexts.¹⁴

3.2 Heterogeneous Effects

To further understand the attributes of the students that were most negatively affected by the intervention, we now examine the impact of the treatment across various baseline attributes. This includes year of study, student baseline GPA, hours worked, and gender. The estimates from this analysis are presented in Table 7.

The results indicate that the students who were harmed by the intervention were in the early years of their degree program (Freshmen or Sophomores), had a baseline GPA above the median value of 3.44, and were working more hours than the median student (12 hours). These results are plausible because less experienced students and those with greater non-school time commitment are more likely to be overconfident and less likely to avail themselves

¹⁴As with most well-designed experiments, our findings have a high degree of internal validity. However, since our experiment focuses on overambitious students, we also encourage future studies to examine academic pledges in other contexts where these adverse effects might not exist.

of campus resources after an initial setback (Nowell and Alston, 2007). The adverse effects of the intervention on higher ability students suggest that these students are more sensitive to the treatment and the negative feedback they received after the first mid-term exam. However, we find no differences in the treatment effect across gender.

4 Discussion and Conclusion

This study uses a field experiment to examine the impact of non-binding commitment pledges in an academic context. We find that commitment pledges when paired with goal setting had a large negative impact on the test scores of students. Our findings also suggest that Freshmen and Sophomores, higher ability, and the students who worked more than 12 hours were most harmed by the treatment, while the treated students in the opposing sub-groups were unaffected.

While this result is surprising, it is consistent with some recent studies in the pledge and goal-setting literature (Sutan et al., 2018; van Lent and Souverijn, 2020; Oppenheimer and Olivola, 2011). For instance, van Lent and Souverijn (2020) found that when students are nudged to target higher goals, it led to a negative effect on their subsequent performance. Similarly, we found that when treated students were asked to pledge the target grade they would work towards in the class, the share of students who increased their targeted grade to an A- or A increased by 11 percentage points. As such, the treated students were setting very ambitious goals and these goals were very difficult to attain in this challenging course.

Unlike other context, in the academic setting, a student who makes a promise to pursue a particular grade has limited control over the attainment of their pursuit. That is, there are factors outside of the students' control that may cause their promised outcome to be unattainable despite their best efforts. As such, since an individual seeks to keep their promise in general (guilt-aversion and preference for promise keeping), a commitment pledge may lock overconfident students into pursuing an unattainable goal. However, as

uncertainty about the final outcome dissipates, the student may become discouraged and adversely change their behavior. We found that while a commitment pledge had no impact on students performance on their first midterm exam, there is a large negative impact on the subsequent performance of treated students. We find strong evidence that the adverse effects we observe are being driven by the disappointment treated students experienced after failing to meet their targeted performance on the first assessment and the subsequent negative effects were strongest for the most overconfident treated students – those with the most unreachable goals.

Our findings suggest that careful attention must be paid while using commitment pledges in the education setting as it may lock students into an overly ambitious goal they are unable to attain and harm their subsequent performance.

References

- Andreoni, J. and Serra-Garcia, M. (2021a). The pledging puzzle: How can revocable promises increase charitable giving? *Management Science*, 67(10):6198–6210.
- Andreoni, J. and Serra-Garcia, M. (2021b). Time inconsistent charitable giving. *Journal of Public Economics*, 198:104391.
- Andrews, R. J., Imberman, S. A., and Lovenheim, M. F. (2020). Recruiting and supporting low-income, high-achieving students at flagship universities. *Economics of Education Review*, 74:101923.
- Ariely, D. and Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological science*, 13(3):219–224.
- Balart, P., Ezquerra, L., and Hernandez-Arenaz, I. (2022). Framing effects on risk-taking behavior: evidence from a field experiment in multiple-choice tests. *Experimental Economics*, 25(4):1268–1297.

- Bearman, P. S. and Brückner, H. (2001). Promising the future: Virginity pledges and first intercourse. *American journal of Sociology*, 106(4):859–912.
- Beattie, G., Laliberté, J.-W. P., Michaud-Leclerc, C., and Oreopoulos, P. (2019). What sets college thrivers and divers apart? a contrast in study habits, attitudes, and mental health. *Economics letters*, 178:50–53.
- Beattie, G., Laliberté, J.-W. P., and Oreopoulos, P. (2018). Thrivers and divers: Using non-academic measures to predict college success and failure. *Economics of Education Review*, 62:170–182.
- Bettinger, E. P. and Baker, R. B. (2014). The effects of student coaching: An evaluation of a randomized experiment in student advising. *Educational Evaluation and Policy Analysis*, 36(1):3–19.
- Bhanot, S. P. (2017). Cheap promises: Evidence from loan repayment pledges in an online experiment. *Journal of Economic Behavior & Organization*, 140:246–266.
- Bryan, G., Karlan, D., and Nelson, S. (2010). Commitment devices. *Annu. Rev. Econ.*, 2(1):671–698.
- Capra, C. M., Jiang, B., and Su, Y. (2022). Do pledges lead to more volunteering? an experimental study. *Economic Inquiry*, 60(1):87–100.
- Castleman, B. L. and Meyer, K. E. (2020). Can text message nudges improve academic outcomes in college? evidence from a west virginia initiative. *The Review of Higher Education*, 43(4):1125–1165.
- Charness, G. and Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, 74(6):1579–1601.
- Clark, D., Gill, D., Prowse, V., and Rush, M. (2020). Using goals to motivate college

- students: Theory and evidence from field experiments. *Review of Economics and Statistics*, 102(4):648–663.
- Clotfelter, C. T., Hemelt, S. W., and Ladd, H. F. (2018). Multifaceted aid for low-income students and college outcomes: Evidence from north carolina. *Economic Inquiry*, 56(1):278–303.
- Cook, P. J., Dodge, K., Farkas, G., Fryer, R. G., Guryan, J., Ludwig, J., Mayer, S., Pollack, H., Steinberg, L., et al. (2014). The (surprising) efficacy of academic and behavioral intervention with disadvantaged youth: Results from a randomized experiment in chicago. Technical report, National Bureau of Economic Research.
- Cotterill, S., John, P., and Richardson, L. (2013). The impact of a pledge request and the promise of publicity: a randomized controlled trial of charitable donations. *Social Science Quarterly*, 94(1):200–216.
- Damgaard, M. T. and Nielsen, H. S. (2018). Nudging in education. *Economics of Education Review*, 64:313–342.
- De Paola, M. and Scoppa, V. (2011). Frequency of examinations and student achievement in a randomized experiment. *Economics of Education Review*, 30(6):1416–1429.
- Dobronyi, C. R., Oreopoulos, P., and Petronijevic, U. (2019). Goal setting, academic reminders, and college success: A large-scale field experiment. *Journal of Research on Educational Effectiveness*, 12(1):38–66.
- Fosgaard, T. and Soetevent, A. R. (2018). Promises undone: How committed pledges impact donations to charity. *Tinbergen Institute Discussion Paper 2018/044-VII*.
- Fryer Jr, R. G. (2017). The production of human capital in developed countries: Evidence from 196 randomized field experiments. In *Handbook of economic field experiments*, volume 2, pages 95–322. Elsevier.

- Fujiwara, T. and Wantchekon, L. (2013). Can informed public deliberation overcome clientelism? experimental evidence from benin. *American Economic Journal: Applied Economics*, 5(4):241–55.
- Goldhaber, D., Long, M. C., Gratz, T., and Rooklyn, J. (2020). Pledging to do “good”: An early commitment pledge program, college scholarships, and high school outcomes in washington state. *Educational Evaluation and Policy Analysis*, 42(1):110–133.
- Hanushek, E. (2010). Education production functions: Evidence from developed countries. *Economics of education*, pages 132–136.
- Hardt, D., Nagler, M., and Rincke, J. (2023). Tutoring in (online) higher education: Experimental evidence. *Economics of Education Review*, 92:102350.
- Himmler, O., Jäckle, R., and Weinschenk, P. (2019). Soft commitments, reminders, and academic performance. *American Economic Journal: Applied Economics*, 11(2):114–142.
- Kajitani, S., Morimoto, K., and Suzuki, S. (2020). Information feedback in relative grading: Evidence from a field experiment. *PloS one*, 15(4):e0231548.
- Kettle, S., Hernandez, M., Sanders, M., Hauser, O., and Ruda, S. (2017). Failure to captcha attention: Null results from an honesty priming experiment in guatemala. *Behavioral Sciences*, 7(2):28.
- Koch, A. K. and Nafziger, J. (2011). Self-regulation through goal setting. *Scandinavian Journal of Economics*, 113(1):212–227.
- Koch, A. K. and Nafziger, J. (2016). Goals and bracketing under mental accounting. *Journal of Economic Theory*, 162:305–351.
- Koch, A. K. and Nafziger, J. (2020). Motivational goal bracketing: An experiment. *Journal of Economic Theory*, 185:104949.

- Krawczyk, M. et al. (2011). Framing in the field. a simple experiment on the reflection effect. *University of Warsaw Faculty of Economic Science Working Paper*, 14:54.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2):443–478.
- Lindo, J. M., Sanders, N. J., and Oreopoulos, P. (2010). Ability, gender, and performance standards: Evidence from academic probation. *American Economic Journal: Applied Economics*, 2(2):95–117.
- McEvoy, D. M. et al. (2016). Loss aversion and student achievement. *Economics Bulletin*, 36(3):1762–1770.
- Meyer, C. J. and Tripodi, E. (2021). Image concerns in pledges to give blood: Evidence from a field experiment. *Journal of Economic Psychology*, 87:102434.
- Morisano, D., Hirsh, J. B., Peterson, J. B., Pihl, R. O., and Shore, B. M. (2010). Setting, elaborating, and reflecting on personal goals improves academic performance. *Journal of applied psychology*, 95(2):255.
- National Center for Education Statistics (2021). Digest of education statistics. Retrieved from https://nces.ed.gov/programs/digest/d20/tables/dt20_226.30.asp.
- Nowell, C. and Alston, R. M. (2007). I thought i got an a! overconfidence across the economics curriculum. *The Journal of Economic Education*, 38(2):131–142.
- Oppenheimer, D. M. and Olivola, C. Y. (2011). *The science of giving: Experimental approaches to the study of charity*. Psychology Press.
- Ordóñez, L. D., Schweitzer, M. E., Galinsky, A. D., and Bazerman, M. H. (2009). Goals gone wild: The systematic side effects of overprescribing goal setting. *Academy of Management Perspectives*, 23(1):6–16.

- Oreopoulos, P., Patterson, R. W., Petronijevic, U., and Pope, N. G. (2022). Low-touch attempts to improve time management among traditional and online college students. *Journal of Human Resources*, 57(1):1–43.
- Page, L. C., Kehoe, S. S., Castleman, B. L., and Sahadewo, G. A. (2019). More than dollars for scholars the impact of the dell scholars program on college access, persistence, and degree attainment. *Journal of Human Resources*, 54(3):683–725.
- Patterson, R. W. (2018). Can behavioral tools improve online student outcomes? experimental evidence from a massive open online course. *Journal of Economic Behavior & Organization*, 153:293–321.
- 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.
- Schippers, M. C., Scheepers, A. W., and Peterson, J. B. (2015). A scalable goal-setting intervention closes both the gender and ethnic minority achievement gap. *Palgrave Communications*, 1(1):1–12.
- Stinebrickner, R. and Stinebrickner, T. (2014). Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model. *Journal of Labor Economics*, 32(3):601–644.
- Stone, N. J. (2000). Exploring the relationship between calibration and self-regulated learning. *Educational Psychology Review*, 12(4):437–475.
- Sutan, A., Grolleau, G., Mateu, G., and Vranceanu, R. (2018). “facta non verba”: An experiment on pledging and giving. *Journal of Economic Psychology*, 65:1–15.
- Thompson, C. A., Eodice, M., and Tran, P. (2015). Student perceptions of general education requirements at a large public university: No surprises? *The Journal of General Education*, 64(4):278–293.

- Tran, A. and Zeckhauser, R. (2012). Rank as an inherent incentive: Evidence from a field experiment. *Journal of Public Economics*, 96(9-10):645–650.
- Tuckman, B. W. (1998). Using tests as an incentive to motivate procrastinators to study. *The Journal of Experimental Education*, 66(2):141–147.
- van Lent, M. and Souverijn, M. (2020). Goal setting and raising the bar: A field experiment. *Journal of Behavioral and Experimental Economics*, 87:101570.
- Vanberg, C. (2008). Why do people keep their promises? an experimental test of two explanations. *Econometrica*, 76(6):1467–1480.
- Wright, N. A. (2020). Perform better, or else: Academic probation, public praise, and students decision-making. *Labour Economics*, 62:101773.
- Wright, N. A. and Arora, P. (2022). A for effort: Incomplete information and college students’ academic performance. *Economics of Education Review*, 88:102238.
- Wrosch, C., Scheier, M. F., Miller, G. E., Schulz, R., and Carver, C. S. (2003). Adaptive self-regulation of unattainable goals: Goal disengagement, goal reengagement, and subjective well-being. *Personality and social psychology bulletin*, 29(12):1494–1508.

Appendix: Main Figures & Tables

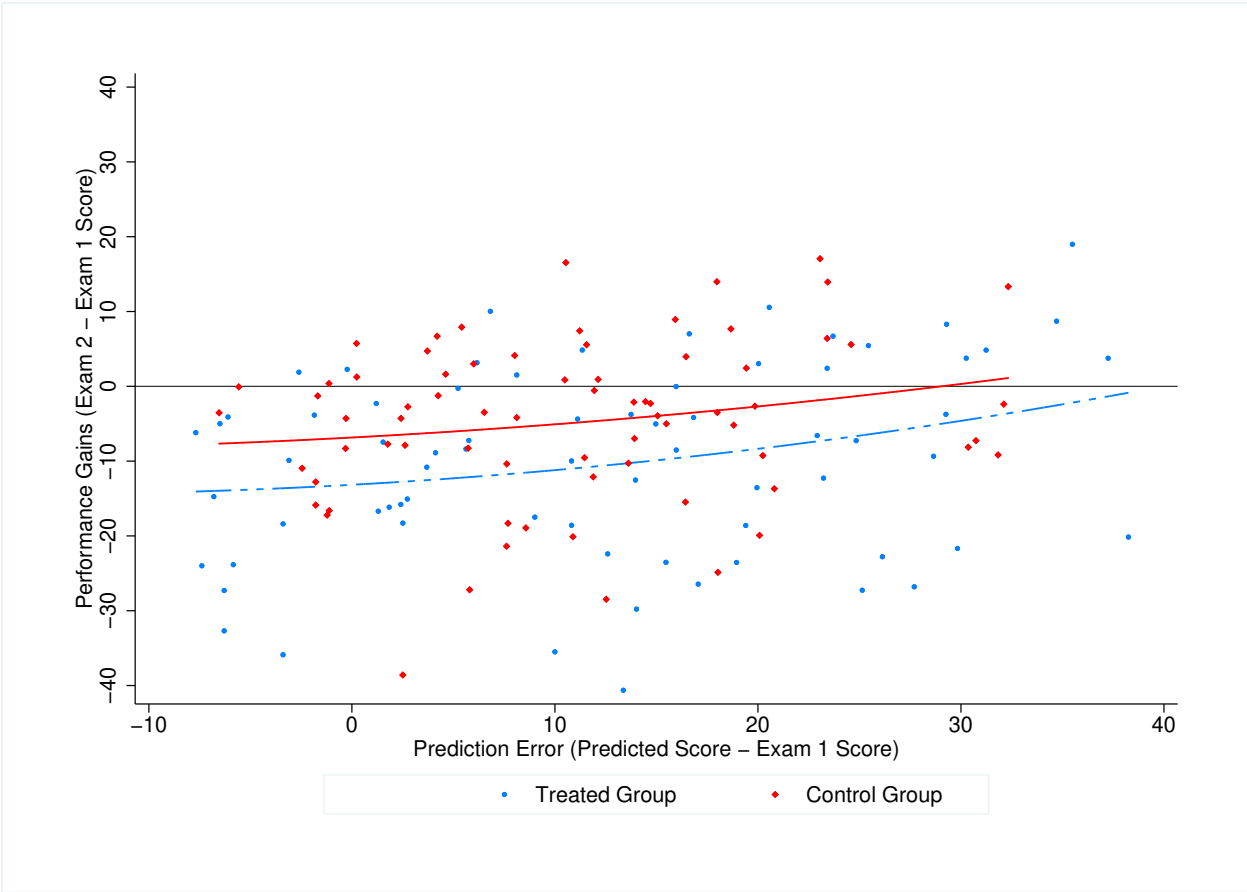


Figure 1: Mid-Term One Prediction Error and Performance Gains on Mid-Term Two

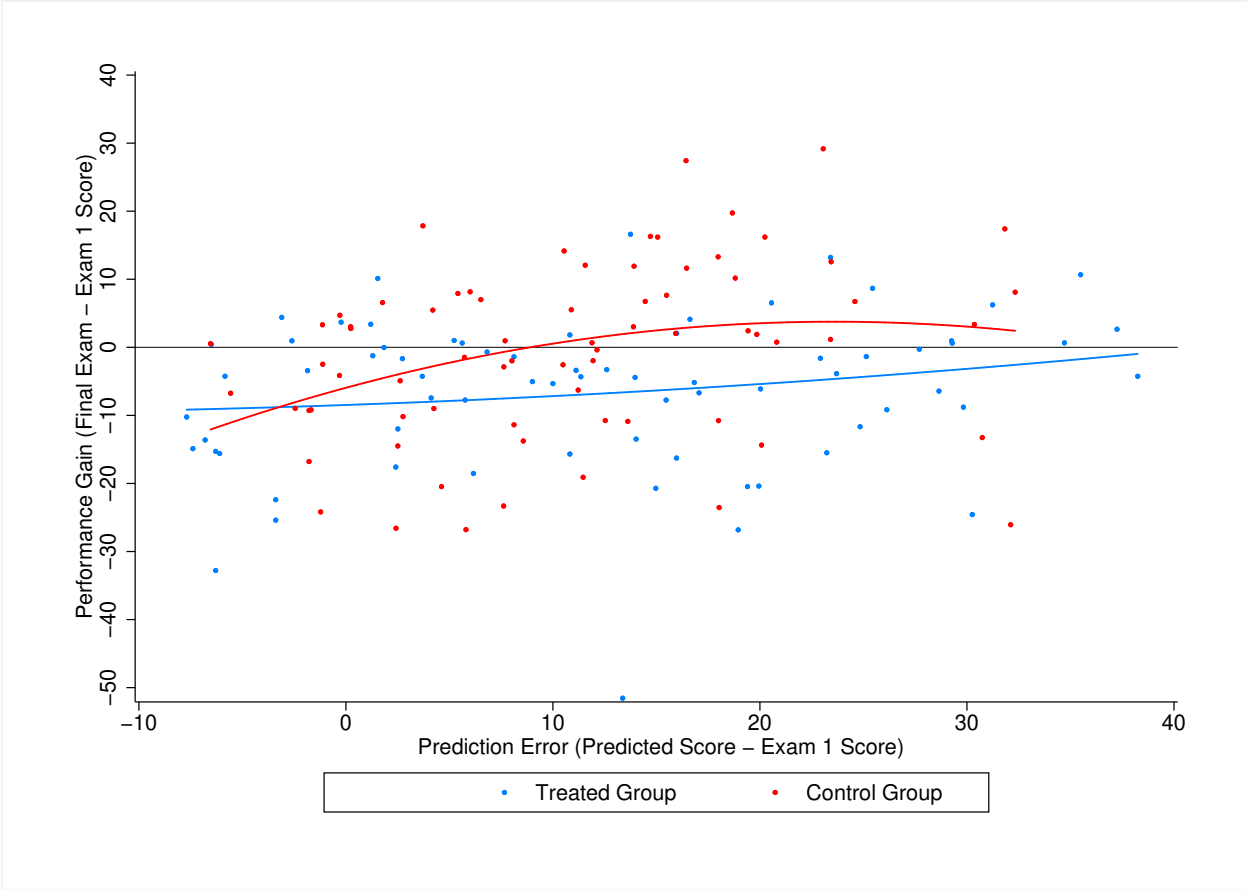


Figure 2: Mid-Term One Prediction Error and Performance Gains on the Final Exam

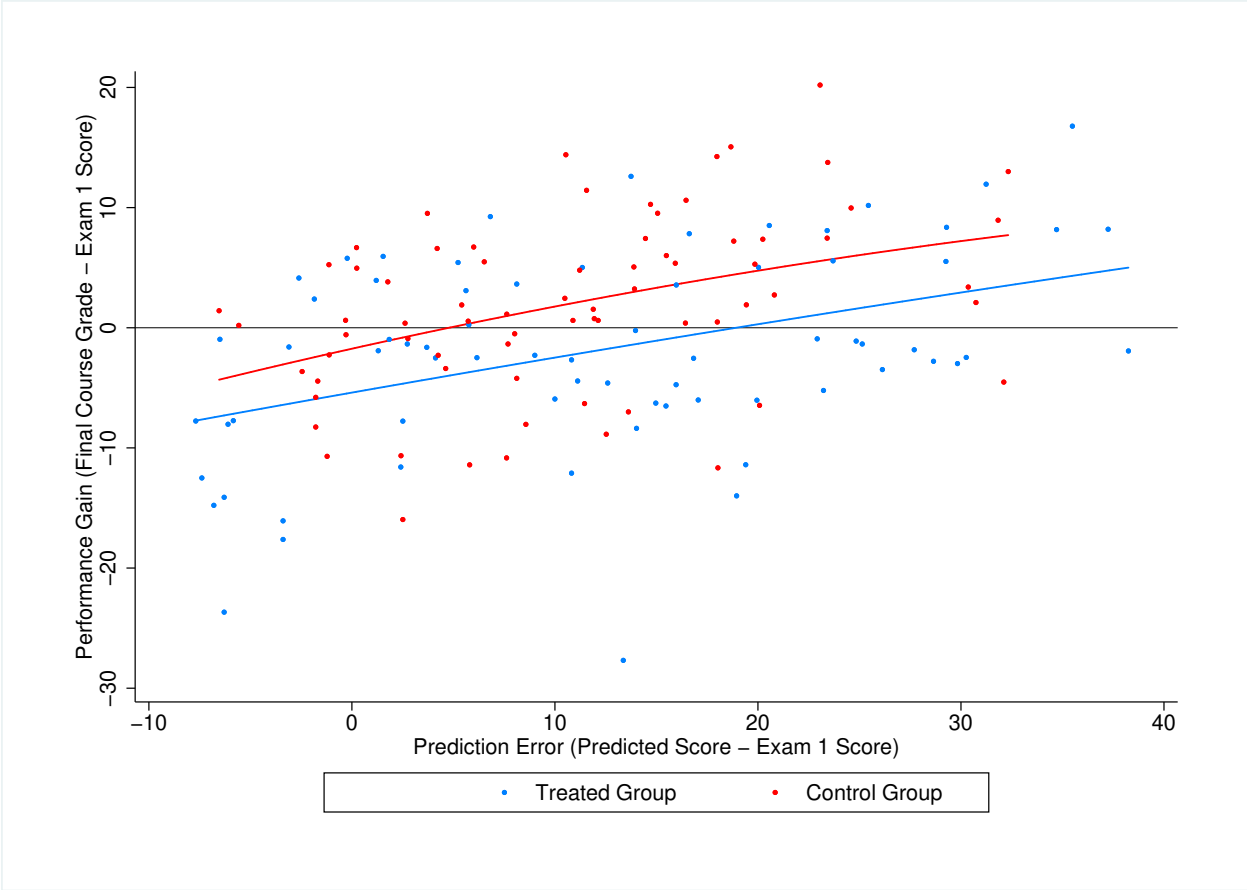


Figure 3: Mid-Term One Prediction Error and Performance Gains in Course Score

	Overall Sample	Control Group	Treatment Group	T-C Diff
Age	20.05 (2.62)	19.83 (2.40)	20.27 (2.82)	0.35 (0.39)
Male	0.63 (0.48)	0.67 (0.47)	0.60 (0.49)	-0.05 (0.07)
Passed Math Pre-Test	0.60 (0.49)	0.57 (0.50)	0.64 (0.48)	0.06 (0.08)
GPA	3.35 (0.56)	3.40 (0.52)	3.29 (0.59)	-0.10 (0.10)
Usual Study Hours	6.52 (3.20)	6.36 (3.13)	6.69 (3.29)	0.30 (0.51)
Number of Courses	4.51 (0.79)	4.56 (0.86)	4.45 (0.73)	-0.11 (0.13)
First Generation	0.21 (0.41)	0.20 (0.40)	0.23 (0.42)	0.02 (0.06)
Freshman or Sophomore	0.69 (0.46)	0.74 (0.44)	0.64 (0.48)	-0.08 (0.08)
Employed	0.66 (0.48)	0.68 (0.47)	0.64 (0.48)	-0.04 (0.07)
Hours Worked	13.46 (13.10)	13.31 (12.33)	13.63 (13.93)	0.09 (2.10)
HH Income \leq 60K	0.43 (0.50)	0.44 (0.50)	0.42 (0.50)	-0.03 (0.08)
Exam 1	73.75 (13.95)	73.73 (14.02)	73.78 (13.97)	-0.05 (2.19)
Exam 2	63.79 (20.99)	65.67 (21.29)	61.83 (20.62)	-3.84 (3.27)
Final Exam	65.63 (24.24)	67.65 (24.90)	63.50 (23.49)	-4.15 (3.78)
Course Score	71.02 (17.48)	72.19 (18.21)	69.80 (16.70)	-2.40 (2.73)
Pass	0.68 (0.47)	0.74 (0.44)	0.63 (0.49)	-0.12* (0.06)
Obs.	164	84	80	164

Table 1: Summary Statistics and Balance Test

Columns 1-3 shows the mean and standard deviation of various background characteristics of students in the overall sample, control group, and treatment group. Column 4 shows the estimated mean difference across the control and treatment group for each background variable. Each equation includes section and instructor fixed effects.

	Predicted Grade	Pledged Grade	Actual Grade
A- or A	0.009 (0.07)	0.11 (0.06)	-0.07** (0.03)
B-, B, or B+	-0.007 (0.06)	-0.09 (0.06)	-0.06** (0.03)
C+ or C	-0.002 (0.01)	-0.02 (0.01)	0.01 (0.01)
D+, D, or Fail			0.11** (0.05)

Table 2: Ordered Probit of Predicted and Actual Letter

Note: Robust standard errors are presented in round brackets below each estimate. Each model includes various pre-treatment covariates: gender, age, expected work hours, math score, gpa, income, instructor and section fixed effects. The dependent variables in columns 1 to 3 are baseline predicted grade, the pledge grades of treated students relative to the benchmark predicted grade in the control group, and actual grades. Level of statistical significance: ***p<0.01, **p<0.05, *p<0.10.

	Mid-Term	Final Exam	Course Grade	Pass
Treated	-0.14 (0.13) [0.29]	-0.28** (0.11) [0.02]	-0.23** (0.11) [0.04]	-0.15** (0.07) [0.03]
Mean	67.64	65.63	71.02	0.68
SD	15.44	24.24	17.48	0.47
Obs	164	164	164	164

Table 3: Main Results, Impact of the Intervention

Note: Robust standard errors and the randomization inference p-values are presented in round and square brackets below each estimate. Each model includes several pre-treatment covariates: gender, age, expected work hours, math score, gpa, income, instructor and section fixed effects. The outcome variables in columns 1-4 are standardized (std) mid-term score, std final exam score, std final course grade, and a pass/fail indicator variable. Level of statistical significance: ***p<0.01, **p<0.05, *p<0.10.

	Mid-Term 1	Mid-Term 2	Homework	% Miss HW	Withdraw
Treated	-0.04 (0.15) [0.81]	-0.27** (0.11) [0.02]	-0.28** (0.11) [0.01]	0.04 (0.03) [0.11]	0.04 (0.04) [0.38]
Mean	73.75	63.79	79.55	0.11	0.09
SD	13.95	20.99	25.91	0.26	0.29
Obs	164	164	164	164	180

Table 4: Impact on Midterms, Coursework, & Dropout

Note: Robust standard errors and the randomization inference p-values are presented in round and square brackets below each estimate. Each model includes various pre-treatment covariates: gender, age, expected work hours, math score, gpa, income, instructor and section fixed effects. The dependent variables in columns 1 to 5 are standardized (std) mid-term score, std final exam score, std overall course grade, and a pass/fail indicator variable. Level of statistical significance: ***p<0.01, **p<0.05, *p<0.10.

	Overperform or Met	Slightly Underperform	Underperform	Grossly Underperform
Mid-Term 2	-2.95 (4.35)	-8.75*** (2.64)	-8.70** (4.03)	1.45 (4.68)
Final Exam	3.78 (3.41)	-10.15*** (3.58)	-12.00** (14.46)	-5.88 (7.28)
Course Grade	-0.11 (2.20)	-4.54*** (1.65)	-6.10** (2.38)	-1.98 (3.64)
Obs	41	41	41	41
Prediction Error	[-7.7, 2.74]	(2.74, 12.04]	(12.04, 20.94]	(20.94, 53.55]

Table 5: Interaction Effect, Treatment and Prediction Error

Note: Robust standard errors and the randomization inference p-values are presented in round and square brackets below each estimate. Each model includes several pre-treatment covariates: gender, age, expected work hours, math score, gpa, income, instructor and section fixed effects. The outcome variables in columns 1-4 are standardized (std) mid-term 2 score, std final exam score, std final course grade, and a pass/fail indicator variable. Level of statistical significance: ***p<0.01, **p<0.05, *p<0.10.

	1	2	3	4	5
Treatment	-0.23*** (0.11)	-0.17** (0.08)	-0.13* (0.07)	-0.07 (0.07)	0.02 (0.05)
Prediction Error		-0.04*** (0.004)	-0.04*** (0.004)	-0.03*** (0.003)	-0.02*** (0.002)
Miss HW			-0.01*** (0.003)	0.01 (0.01)	0.005* 0.003
Homework Score				0.50*** (0.14)	0.35*** (0.07)
Mid-Term 2					0.53*** (0.05)
<i>Indirect Effect</i> <i>Overall Effect</i>		26%	43%	70%	100%

Table 6: Mediation Analysis

Note: Robust standard errors are presented in round and square brackets below each estimate. Each model includes several pre-treatment covariates: gender, age, expected work hours, math score, gpa, income, instructor and section fixed effects. The outcome variables in all columns is the standardized overall course grade. Level of statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. The estimate in the first column is the total treatment effect. The last row shows the share of the total effect that is mediated by the variables included in the model.

	Mid-Term	Final Exam	Course Grade	Pass
Freshmen and Sophomores	-0.34** (0.16) [0.03]	-0.43*** (0.13) [0.00]	-0.41*** (0.13) [0.00]	-0.20** (0.08) [0.03]
Juniors and Seniors	0.003 (0.22) [0.96]	-0.05 (0.23) [0.77]	-0.04 (0.21) [0.75]	-0.02 (0.13) [0.85]
GPA Above Median	-0.37* (0.20) [0.08]	-0.47*** (0.16) [0.00]	-0.44** (0.17) [0.01]	-0.25*** (0.09) [0.00]
GPA Below Median	0.10 (0.14) [0.78]	-0.06 (0.16) [0.96]	0.03 (0.14) [0.94]	-0.03 (0.11) [0.80]
Work Hours above median	-0.06 (0.19) [0.13]	-0.34** (0.16) [0.01]	-0.20 (0.16) [0.04]	-0.20** (0.10) [0.03]
Work Hours below median	-0.30 (0.19) [0.59]	-0.29 (0.18) [0.67]	-.34* (0.17) [0.65]	-0.07 (0.10) [0.77]
Male	-0.24 (0.17) [0.13]	-0.21 (0.14) [0.15]	-0.23 (0.14) [0.08]	-0.08 (0.09) [0.31]
Females	0.05 (0.21) [0.68]	-0.30 (0.18) [0.16]	-0.15 (0.18) [0.47]	-0.19* (0.11) [0.11]

Table 7: Heterogeneity Analysis

Note: Robust standard errors and the randomization inference p-values are presented in round and square brackets below each estimate. Each model includes several pre-treatment characteristics: gender, age, expected work hours, math score, gpa, income, instructor and section fixed effects. The outcome variables in columns 1-4 are standardized (std) mid-term score, std final exam score, std final course grade, and a pass/fail indicator variable. Level of statistical significance: ***p<0.01, **p<0.05, *p<0.10.

Online Appendix A: Additional Details

	Predicted Grade	Pledged Grade	Baseline Time Com.	Revised Time Com.
A- or A	41.25%	52.50%	4.91	8.88
B-, B, or B+	56.25%	46.25%	5.44	9.3
C+ or C	2.50%	1.25%	6	6.7
D+, D, or Fail	0.00%	0.00%		

Table 10: Treated Students Prediction, Goal Setting, & Time Investment

Note: The dependent variables in columns 1-4 are baseline predicted grade, the pledged grades of treated students, time commitment prior to signing the pledge, and the promised time commitment after signing the pledge. These descriptive statistics are only applicable to treated students.

	1	2	3	4	5
Treatment	-0.23*** (0.11)	-0.22** (0.07)	-0.14 (0.11)	-0.02 (0.07)	0.003 (0.05)
Exam 1		0.53*** (0.05)			
Homework Score			0.02*** (0.003)		
Exam 2				0.76*** (0.05)	
Final Exam					0.88*** (0.04)
<i>Indirect Effect</i>		4%	39%	91%	100%
<i>Overall Effect</i>					

Table 20: Mediation Analysis

Note: Robust standard errors are presented in round and square brackets below each estimate. Each model includes several pre-treatment covariates: gender, age, expected work hours, math score, gpa, income, instructor and section fixed effects. The outcome variables in all columns is the standardized overall course grade. Level of statistical significance: ***p<0.01, **p<0.05, *p<0.10. The estimate in the first column is the total treatment effect. The last row shows the share of the total effect that is mediated by the variables included in the model.



Informed Consent

You are invited to participate in a research study conducted through Florida Gulf Coast University. You must be age 18 or older to participate in the study. Your participation in this study is voluntary. The University requires that I obtain your signed consent to participate in this study. Refusal to join the study will not affect any future services you may be eligible to receive from the University.

Study Summary

This study aims to understand how pledges affect students' academic outcomes.

The results from this study will help policymakers devise policies that can improve students' performance and how these college courses are delivered in future semesters. Your responses will be kept confidential and completing the survey poses 'less than minimal' risks to your physical, emotional, social, professional, or financial well-being.

If you are interested in learning more about this study, please continue to read below and use the contact information at the end of this document to ask the researcher any questions you may have.

Purpose of the Study

The purpose of this study is to examine how students form beliefs about their expected performance in college courses.

Invitation to Participate

I am asking you to take part in the study because you can provide meaningful information about how students form beliefs about the courses they take and how these beliefs affect their effort and performance.

Description of your Involvement

All participants are required to complete this survey. The researcher is also requesting your consent to access and utilize your academic records (grades) in the study. The researcher will match your responses on this survey to your in-class performance. Your class grades will be kept confidential.

Benefits of Participation

The results from this study will likely improve how courses are delivered at FGCU. In addition, the study can have a meaningful impact on student success at higher education institutions across the country.

Risks and Discomforts of Participation

Your participation and academic data will be kept confidential. However, working with email or the internet has a small risk of compromising privacy, confidentiality, and/or anonymity. Despite this possibility, the risks to your physical, emotional, social, professional, or financial well-being are considered to be 'less than minimal' by participating in the study.

Compensation for Participation

Extra credit points will be awarded to the entire class based on the level of participation. There is a potential of earning up to 5% on your first exam. This is based on the level of class participation. For example, if 80% of the class completes the survey, then you will receive 4% on your first exam.

Confidentiality

If you join the study, I will make every effort to keep your information confidential and secure by ensuring that (i) your survey responses are kept confidential and (ii) only the principal investigator(s) will have access to any educational data that will be collected (iii) your identifying information is removed from the data as soon as possible. However, despite these safeguards, there is the possibility of hacking or other security breaches that could compromise the confidentiality of the information you provide. Thus, it is important to remember that you are free to opt-out of the study.

I will not release information about you unless you authorize us to do so or unless we are required to do so by law. If the results of this study are published or presented at a professional meeting, no information will be included that would make it possible to identify you as a study participant.

It is possible that organizations responsible for making sure the research is done safely and properly such as the university, and government offices may need to see the information you provide.

Voluntary Nature of Study

Participating in this study is completely voluntary. Refusal to join the study will not affect any future services you may be eligible to receive from the University. You can choose to discontinue your participation at any time by simply closing your web browser and not complete the survey. If you choose to join the study, you can leave it at any time with no penalty.

Contact Information for the Study Team

I do not foresee any medical problems from participating in this study. If you have any questions or concerns about this study, please contact Dr. Nicholas Wright at nawright@fgcu.edu or Dr. Jesse Wright at jwright@fgcu.edu or Dr. Puneet Arora at puneet.arora@ahduni.edu.in.

Statement: I have read the preceding information describing this study. All of my questions have been answered to my satisfaction. I am 18 years of age or older and freely consent to participate in the study. My decision to participate or to decline participation in this study is completely voluntary. I understand that I am free to withdraw from the study at any time. I am aware of my option to not answer any questions I choose.

I understand that it is not possible to identify all potential risks I believe that reasonable steps have been taken to minimize both the known and potential but unknown risks. The submission of the completed survey is my informed consent to participate in the study. If you would like a copy of the consent form, print a copy before continuing.

- I consent
- I do not wish to participate (opting out of the survey will lower the extra credit for everyone)

Socioeconomic Questions

What is your assigned sex at birth?

- Male
- Female

What is your month and year of birth?

Year of Birth

Month of Birth

Are you Spanish, Hispanic, or Latino or none of these (select all that apply).

- Spanish
- Hispanic
- Latino
- None of these

Choose one or more races that you consider yourself to be:

- White
- Black or African American
- American Indian or Alaska Native
- Asian
- Native Hawaiian or Pacific Islander
- Other (specify)

To which college do you belong?

- Arts and Sciences
- Lutgert College of Business
- Marieb College of Health & Human Services
- Education
- Whitaker College of Engineering

What is your major?

Which year do you belong to?

- Freshman
- Sophomore
- Junior
- Senior

Information about income is very important to understand. Would you please give your best guess?

Please indicate the answer that includes your entire household income in (previous year) before taxes.

- Less than \$10,000
- \$10,000 to \$19,999
- \$20,000 to \$29,999
- \$30,000 to \$39,999
- \$40,000 to \$49,999
- \$50,000 to \$59,999
- \$60,000 to \$69,999
- \$70,000 to \$79,999
- \$80,000 to \$89,999
- \$90,000 to \$99,999
- \$100,000 to \$149,999
- \$150,000 or more

Are you the first in your immediate family to attend college?

- Yes
- No

Did either of your parents or guardians complete college?

- Yes
- No

Where did you rank in your graduating high school class?

- Top 5%
- Top 10%
- Top 15%
- Top 20%
- Top 25%
- Top 40%
- Top 50%
- Bottom 50%

How would you rate your typical academic performance?

- Excellent
- Above Average
- Average
- Below Average
- Poor

What is your combined SAT Score? Insert 0 if you have not taken the exam.

0 160 320 480 640 800 960 1120 1280 1440 1600

0 160 320 480 640 800 960 1120 1280 1440 1600

SAT Score

What is your ACT Score? Choose 0 if you have not taken the exam.

0 3 6 9 12 15 18 21 24 27 30 33 36

ACT Score

What is your current GPA (4.0 Scale).

0 1 2 3 4

GPA

In a typical week, how many hours did you spend reading, studying, completing assignments, and preparing for your most challenging course last semester?

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

Hours

How challenging do you think this course will be?

- Very easy
- Somewhat easy
- Neither easy nor difficult
- Somewhat difficult
- Very difficult

Risk Assessment

Please indicate which of the following game show options you would prefer: option A or option B.

Option A	Option B
A lottery with a 10% chance of receiving \$2 and a 90% chance of receiving \$1.60	A lottery with a 10% chance of receiving \$3.85 and a 90% chance of receiving \$0.10

- Option A
 Option B

Please indicate which of the following game show options you would prefer: option A or option B.

Option A	Option B
A lottery with a 20% chance of receiving \$2 and a 80% chance of receiving \$1.60	A lottery with a 20% chance of receiving \$3.85 and a 80% chance of receiving \$0.10

- Option A
 Option B

Please indicate which of the following game show options you would prefer: option A or option B.

Option A	Option B
A lottery with a 30% chance of receiving \$2 and a 70% chance of receiving \$1.60	A lottery with a 30% chance of receiving \$3.85 and a 70% chance of receiving \$0.10

- Option A
 Option B

Please indicate which of the following game show options you would prefer: option A or option B.

Option A	Option B
A lottery with a 40% chance of receiving \$2 and a 60% chance of receiving \$1.60	A lottery with a 40% chance of receiving \$3.85 and a 60% chance of receiving \$0.10

- Option A
 Option B

Please indicate which of the following game show options you would prefer: option A or option B.

Option A	Option B
A lottery with a 50% chance of receiving \$2 and a 50% chance of receiving \$1.60	A lottery with a 50% chance of receiving \$3.85 and a 50% chance of receiving \$0.10

- Option A
- Option B

Please indicate which of the following game show options you would prefer: option A or option B.

Option A	Option B
A lottery with a 60% chance of receiving \$2 and a 40% chance of receiving \$1.60	A lottery with a 60% chance of receiving \$3.85 and a 40% chance of receiving \$0.10

- Option A
- Option B

Please indicate which of the following game show options you would prefer: option A or option B.

Option A	Option B
A lottery with a 70% chance of receiving \$2 and a 30% chance of receiving \$1.60	A lottery with a 70% chance of receiving \$3.85 and a 30% chance of receiving \$0.10

- Option A
- Option B

Please indicate which of the following game show options you would prefer: option A or option B.

Option A	Option B
A lottery with a 80% chance of receiving \$2 and a 20% chance of receiving \$1.60	A lottery with a 80% chance of receiving \$3.85 and a 20% chance of receiving \$0.10

- Option A
- Option B

Please indicate which of the following game show options you would prefer: option A or option B.

Option A	Option B
A lottery with a 90% chance of receiving \$2	A lottery with a 90% chance of receiving \$3.85

and a 10% chance of receiving \$1.60

and a 10% chance of receiving \$0.10

- Option A
- Option B

Please indicate which of the following game show options you would prefer: option A or option B.

Option A	Option B
A lottery with a 100% chance of receiving \$2 and a 0% chance of receiving \$1.60	A lottery with a 100% chance of receiving \$3.85 and a 0% chance of receiving \$0.10

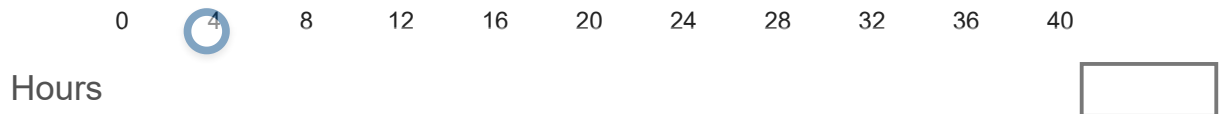
- Option A
- Option B

Time Use and Course Information

How many courses are you taking this semester?



How many **hours** do you work **each week** on average? (Include 0 if you do not work)



Did you talk to your friends about this course before enrolling?

- Yes
- No

Did you get information about the instructor before enrolling in this course? (from rate my professor, friends, etc)

- Yes
 No

What grade do you predict you will earn in this class?

- A** (I will score between 92.5% & 100%)
 A- (I will score between 89.50% & 92.49%)
 B+ (I will score between 86.5% & 89.49%)
 B (I will score between 82.50% & 86.49%)
 B- (I will score between 79.50% & 82.49%)
 C+ (I will score between 76.50% & 79.49%)
 C (I will score between 69.50% & 76.49%)
 D (I will score between 59.50% & 69.49%)
 Fail (I will score below 59.50%)

To get the grade you expect for this course, how many hours will you devote to reading, studying, completing assignments, and preparing for **this class** each week?

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Hours

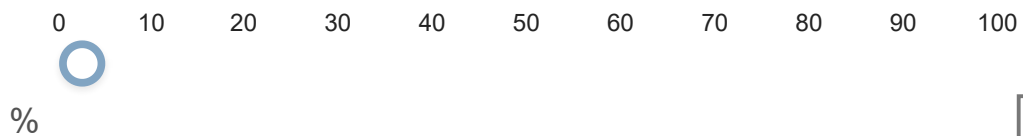
Treatment

A pledge to my professor and myself: I promise to attend the Principles of Macroeconomics classes regularly, complete and

submit all assignments on time, study hard, and appear for the exams with utmost sincerity. I pledge that I will take these action to obtain a target grade of:

- » **A** (I will score between 92.5% & 100%)
- » **A-** (I will score between 89.50% & 92.49%)
- » **B+** (I will score between 86.5% & 89.49%)
- » **B** (I will score between 82.50% & 86.49%)
- » **B-** (I will score between 79.50% & 82.49%)
- » **C+** (I will score between 76.50% & 79.49%)
- » **C** (I will score between 69.50% & 76.49%)
- » **D** (I will score between 59.50% & 69.49%)
- » **Fail** (I will score below 59.50%)

To obtain your target grade, what percent of classes do you promise to attend?



To obtain your target grade, how many hours do you promise to devote to the quizzes and homework assignments in this class each week?



To obtain your target grade, how many hours do you promise to devote to reading for this class each week?

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Hours

Sharing Pledge

Sharing your pledge with a close friend or relative may help you to stay on track to achieve your target grade. Do you want us to share your pledge with a friend, relative, or your academic advisor? If yes is selected, we will only share your pledge and no other academic information.

- Yes
 No

Pledge Recipient

Please provide the contact information for a friend or relative that is **not** currently taking this course:

Email Address:

Cellphone Number:

Relation to recipient: