

Less Work, More Learning?

Susan Jacobson
Regis College

This paper uses the ideas of experimental economics to see if a change in grading structure increases motivation to study. The treatment group of students were rewarded for strong performance on a first exam by being excused from a second exam. The tentative results indicate that perhaps the reward structure did motivate students to study harder for the first exam. But there is also some evidence of unintended consequences, that those students who did not get automatic As on the second exam were discouraged and less motivated than they otherwise would have been to study for the second exam.

SETTING THE STAGE

My biggest challenge in teaching the required course in business finance to students who are predominantly marketing and management majors is motivating them to meet the challenge of the quantitative demands necessary to master the basic principles of the course. In the last few years, about one-third of the students in these classes have had to withdraw because they are failing. This is difficult to understand because at an even mildly selective university, the vast majority of the students must have had the necessary training to do the math required for basic finance. After all, the most difficult calculations are generally similar to a beginning algebra class. For this reason, failure in this class seems to me to be primarily an issue of being motivated to put in the necessary number of hours of study. The intuition, supported by observing classroom behavior, is that this is more a problem of motivation rather than ability and/or preparation. Perfect scores on exams by some students support this intuition.

For decades economists have used highly mathematical models to derive economic theory that explains decision making. These theories are all dependent on a large number of assumptions which are most often abstractions from reality and rarely tested. Because economics depends heavily on what motivates people to choose one thing over another, many economists are now using a scientific, controlled-experiment approach to look for solutions to problems. *The Why Axis* by Uri Gneezy and John A. List (2013) reports how this method can be used to answer questions about gender discrimination, underperformance in public schools, charitable giving, and creativity in business. The authors' message is, "If you want people to do what you want, incentives can be incredibly handy" (p. 18).

Because underperformance in business finance seems to be somewhat endemic, I was particularly taken by Gneezy and List's experiments to determine what motivates students to perform. I have tried a variety of methods to motivate students: daily quizzes, surprise quizzes, participation grades, homework assignments, homework assignments collected at random, homework solved on the board by students – the list seems to go on and on, but nothing has worked as well as hoped. What's the answer? Gneezy and List (2013) write, "So how do we learn about underlying motivations and the right incentives?" (p. 3). Their answer is experimentation. "Fortunately, field experiments can provide the kind of hard data that

citizens, educators, philanthropists, policy makers, and CEOs need to not only avoid making mistakes but also to develop a better understanding of the people they are supposed to serve: What really motivates people and why?" (Gneezy & List, 2013, p. 9). Importantly, we all must figure out what it is that might prompt people to make the kinds of decisions we wish they would – in the case of business finance, it is to study harder. As Gneezy and List (2013) say, "Here is the truth. If you want people to do something, you really need to understand what motivates them" (p. 31). Because they found a number of things that worked to motivate public school students in Chicago, reading Gneezy and List prompted me to think about what might motivate Regis College students by getting into *their* heads, not trying to reason from my theoretical reality. From observing behavior, I concluded students would be motivated by a reward that reduced their perceived work load. Of course the intent is NOT to reduce learning, but to increase motivation and learning by "making it easier" in perception, but not in reality.

Gneezy and List (2013) conclude, "Our explorations into public education have taught us the power of combining field experiments with economic reasoning" (p. 89). That economic reasoning is the fundamental neoclassical assumption that people maximize utility. Once you figure out what creates utility for people, you're on to the secret of motivation. That's why I began to try to think like a student instead of a professor. The experiment in Business Finance attempted to increase student utility by rewarding them with a reduced work load for greater effort early in the course, with the hope that this would also result in greater learning over the entire semester.

THE EXPERIMENT

Following Gneezy and List's lead, I experimented with exam requirements in business finance. Students who studied harder in the first weeks of the semester, the weeks which are absolutely critical to success in the course, were rewarded with less work – an opportunity to skip the second exam.

Two sections of business finance in spring 2014 provided the small sample for this experiment. We used the text *Principles of Finance with Microsoft Excel* by Simon Benninga and did spreadsheet work in class at least 70% of the time. Students self-selected into one or the other section during the registration process. In the later section (10:50 AM to 12:05 PM) the students were required to take two exams during the semester and a comprehensive final at the end of the semester. This section was the control group – their exam schedule followed my usual custom. The earlier section (9:25 AM to 10:40 AM) served as the treatment group. These students were rewarded for high scores on the first exam. Any student who scored an A- or above on the first exam was allowed to skip the second exam and be awarded the same grade as that on the first exam for the second exam. This first exam is critical because it lays the groundwork, time value of money, for much of the rest of the material in the class – bond valuation, stock valuation, and capital budgeting. Anyone who masters this first chunk of material finds the rest of the material quite accessible.

Students in the first section who scored Bs or Cs on the first exam had a choice whether or not to take the second exam or just accept the grade from the first exam for their second exam. Essentially, they were offered the opportunity to improve their grades by taking a second exam. Anyone scoring less than a C- on the first exam was required to take the second exam. Anyone scoring below a C- on both exams was advised to withdraw from the class. (Students who do not earn at least a C- in the class must take it over anyway.) Everyone was required to take the final exam. Because the final is comprehensive, this ensured that all students who remained in the class had an incentive to master all the material at least at the C- level.

The intent was to see if those in the first section were more motivated to really learn the crucial material in the beginning of the course. Allowing students to skip the second exam was intended to "reward" these students with a reduced work load – that was the carrot. On the other hand, the clear penalty, articulated on the syllabus, that they would have to take a second exam and be advised to withdraw if they did not master at least enough of the material to get a C- on both exams, was the stick. (As a rule, because it is virtually impossible to gain even a basic competency after performing that badly on the first two exams, students who score below a C- on the first two exams have always been advised to

withdraw from the class. But this policy has not been stated clearly on the syllabus as it was for the first section.)

This experiment, approved by the Regis University IRB, posed no "risk" to any student. If a student in the treatment group wanted to be a part of the control group, that student could simply choose to take all three exams and essentially replicate the conditions of the control group (grading as usual). If any student had expressed concern about being in the treatment group, that student would have been counseled to take all three exams and be graded as the students in the control group were graded. Interestingly, no student in either section complained about the grading in her/his section.

THE RESULTS

These results are simply suggestive given the small sample sizes and the one-time experiment. But they are interesting enough to consider and discuss.

**TABLE 1
DIFFERENCES IN PERFORMANCE**

Section	# Students	Exam 1	Exam 2*	Final Exam	Course	GPA*
9:25	21	2.6	2.5	2.0	2.3	2.99
10:50	20	2.5	3.2	2.3	2.7	3.38

*Significantly different at the 1% level.

From the data in Table 1 we can make some conjectures. First, it's important to note that the average GPA, when the students entered the class, was significantly higher in the later section. (Note that these GPAs are considerably higher than the grades for this course.) I can only hypothesize why there is this difference in GPAs. It seems to me that students prefer being in the 10:50 section because it is later. If that's true, students who are paying more attention to when they are able to register and are conscientious about being first in line, get into the second section before it closes. Then many of the students who are defaulted into the later section may be students who are less conscientious about their school work in general. From my observation, the difference between the two sections' abilities and work ethics was apparent and meaningful to the classroom environment in this particular semester.

Despite this difference in GPA, the scores on the first exam were almost identical, though the earlier group did slightly outscore the later section. This supports the hypothesis that the "treatment" worked, that students in the earlier section studied harder than they otherwise would have and earned grades that were higher than expected relative to their overall GPAs.

The results for the second exam, though disappointing for the earlier section, reflect the elimination of the better students from the sample. The three students in the first section who scored As on the first exam, did not take the second exam – the sample is missing the high performers. Also, of the five students in the first section who earned Bs on the first exam, one student opted out of taking the second exam. (This was a student who had just become a first-time father.) Somewhat troubling though is the result that all the other students in the earlier section who earned Bs on the first exam and took the second exam by choice, earned lower grades on their second exams. This contrasts with the higher average grade on the second exam for the second section where 16 of 20, and all those who earned Bs on the first exam, raised their grades from the first exam to the second. These students, in the 10:50 section, may have been motivated by the gap between their performance on the first exam and their expected performance as predicted by their GPAs. The poorer performance in the second section may also reveal that the perceived penalty of having to take the second exam when not all of their classmates had to take it reduced the motivation to study for the second exam. This would be an example of what we see so often from incentive systems – unintended consequences.

In the end, the grades on the comprehensive final and the course grades were not significantly different. Does this mean we can conclude that the experiment worked because the course grades are not significantly different but the GPAs are?

Final course grades are, of course, dependent on a number of variables, the most important perhaps being ability and an already-formed work ethic. (In the last two semesters in six different classes, self-reported hours studied is NOT significantly correlated to self-reported GPA in a cross sectional analysis. This does not mean that any individual could not do better by studying more, but it does indicate that “talent” has a lot to do with grades.) These two characteristics are reflected in the students’ GPAs. Another variable that seems to impact performance is chosen major. Accounting students tend to do better in finance than marketing students because success in finance demands some recall of accounting principles. So I recorded which students were marketing and which were accounting majors. In a multiple regression analysis, GPA and the dummy variable for accounting majors were significant and positive at the 1% level. But the section dummy was insignificant. (Because the GPA in the later section is statistically significantly higher for the second section, the GPA variable and the section variable were not orthogonal as they should be for a proper regression analysis.) The insignificance of the section dummy may indicate a failed experiment – the students in the first section were either not motivated by the grading structure or motivation was insufficient to influence final grades in a significant way. However, if we look at the data in a different way, we might conclude that the grading structure in the first section did provide motivation to perform better than the students would have otherwise because the final and course grades were not significantly different despite the difference in GPAs. Also, because the average scores on the first exam were almost identical, despite higher GPAs in the second section, it appears that perhaps the students in the first section were motivated to study harder for the first exam. Unfortunately performance on the second exam suggests that perhaps there may have been unintended consequences of the incentive structure in the first section.

In addition to looking at performance, I also looked at perceptions. I wanted to find out whether the students thought the way the exams and grading were structured had any impact on their motivation. To assess this, students rated the four statements below on a five-point Likert scale (1 = Strongly disagree, 2 = Disagree, 3 = Neither Agree nor Disagree, 4 = Agree, 5 = Strongly Agree).

1. The option to not have to take the second exam if I did well on the first exam prompted me to work harder to do well on the first exam.
2. The option to improve my grade with the second exam made me work less hard to prepare for the first exam.
3. Having a comprehensive final makes me study harder for the final.
4. Having a comprehensive final makes me study harder for all of the exams.

Unfortunately only 13 of the 21 students were in class the day they filled out the survey. (That in itself tells you something about motivation.) Nevertheless, 10 of the 13 students who did respond to the survey, either agreed or agreed strongly with the first statement – the vast majority of the students indicated that they were motivated to work harder for the first exam. Ten of the 13 students responded to the second question by disagreeing or strongly disagreeing. This is the hoped-for result – that the option to “try again” would not decrease motivation to study for the first exam. Unfortunately though, recall that all the students who earned Bs on the first exam who also took the second exam, had lower grades on the second exam. Not anticipating this result, I did not ask the students whether not scoring high enough on the first exam to essentially get an A on the second exam without taking it resulted in less motivation to study for the second exam. Again, this emphasizes the problem of unintended consequences.

In response to the third statement, 10 students again either agreed or strongly agreed that a comprehensive final makes them study harder. This seems so obvious that it’s curious that the other three did not agree. Disappointingly, only seven of 13 respondents agreed with the final statement. This is unfortunate because finance is a subject that is virtually impossible to cram for right before a comprehensive final.

CONCLUSIONS

We all struggle with ways to get the best out of students. Unfortunately, we seldom have experimental data that show us how to do that. This paper is a first attempt to use the experimental method becoming popular among economists to test an idea to improve performance through incentives. The results, though interesting and somewhat encouraging, also point out the problem of unintended consequences. This problem highlights why we need experimental data. Logical reasoning and theory often miss real consequences. One painful example of this was when stock options, the answer to the agency problem, backfired. Clearly, structuring incentives to achieve desired results is complicated.

REFERENCES

Gneezy, U. & List, J.A. (2013). *The Why Axis*. New York: Public Affairs.