

Giving Feedback to Students in Large Lecture

Kelvin Wong

Arizona State University

Draft*

Abstract

In this paper, I investigate the impact of giving personalized feedback to students via email over the course of the semester. A feedback email typically explains to students exactly how they are doing in the class, and what they can be doing to improve on their performance. All student performance data is collected from Principles of Microeconomics courses at a large state university. First, I give a method for how personalized feedback can be given in a large lecture. Second, I analyze the performance of students in classes where personalized feedback is given after exams and compare it with classes where it is not. I find that students receiving the lowest and highest scores after a first exam are most helped by personalized feedback. Lastly, I discuss the results and give some anecdotal evidence that personalized feedback positively impact student performance.

* *This is a rough draft of a more formal paper I hope to come from this. As such, there is currently no literature review, but simply a bare-bones version that hopefully conveys what the final paper will be about.*

1 Introduction

The large lecture is common-place in higher education, and often students treat these classes as ones they can just go in and out without ever interacting with the professor. Even well-meaning professors struggle to find ways to connect with so many students given time constraints. In a world with no constraints, a professor can meet every student, understands their needs, gives them guidance and encouragement, and ensures that they succeed.¹ However, this is simply not possible in larger classes (and perhaps even in smaller ones). In this paper, I explore a way to connect with students through giving personalized feedback and whether this improves student performance.

Personalized feedback in this paper refers to emails that are sent to students after exams to let them know how they are doing in the class. While students can certainly go to their gradebook and look at their grades over the course of the semester, personalized emails provide recommendations on how to do better, encouragement to visit office hours, and a confirmation of the professor's dedication to each student's learning. The last point may seem insignificant on the surface, but from anecdotal evidence (via emails, office hours, and conversations outside of class) students do think that such emails boost their morale and leads them to work harder.

The semester that I started doing personalized feedback to students in my own classes, I gave students a survey near the end of the semester to assess how they perceived personalized feedback.² In Table 1, I summarize the results from that survey.

Table 1: Survey on Effectiveness of Personalized Feedback

Survey Question	Mean Response (N=153)
Receiving feedback after the first exam was helpful (1: not helpful, 5: very helpful)	4.42
Receiving feedback after the second exam was helpful (1: not helpful, 5: very helpful)	4.50
I changed how I studied after receiving feedback from exam 1 (1: strongly disagree, 5: strongly agree)	3.35
I changed how I studied after receiving feedback from exam 2 (1: strongly disagree, 5: strongly agree)	3.24
I felt more invested in the class after receiving feedback. (1: strongly disagree, 5: strongly agree)	4.10
Getting feedback on how I'm doing is something I expect from every class. (1: strongly disagree, 5: strongly agree)	3.60
Receiving feedback made me work harder. (1: yes, 0: no)	.80
Receiving feedback helped me to do better in this class. (1: yes, 0: no)	.91
I came to office hours because of receiving feedback (and I would not have otherwise) (1: yes, 0: no)	.23

While what students perceive and what is actually true are two different things, I still believe the results of this survey is significant. For instance, students tend to overwhelmingly believe that receiving feedback helped them do better in the class. And while it could have been the feedback that helped, it could also have been their own motivation increasing throughout the semester, or the availability of resources that

¹ However, in such a world, economists also would not have jobs.

² Survey was given to students at a different large state university than the one for which data was collected for this paper.

they utilized. But does the answer to that question really matter? On a research level, perhaps. But in practice, students do believe feedback helps and that may ultimately be more important. If students feel that they are being invested in, they tend to reciprocate by investing more into the course. I also think the increase of students to office hours help tremendously. I had kept track of students who came to office hours because of personalized feedback and tracked their progress in the class, and most of them will see greater improvements than their classmates over the course of the semester. I have not had the time to put this data together, but I will present this data in a future version of the paper.

All this to say, students at the least appreciate receiving proper feedback. It also seems like most students in large lectures do not expect feedback, so getting them may mean more. While I do not have hard evidence for this, I do hear it all the time from students with comments along the lines of “receiving this email from you means that you care, and it makes me want to work so much harder.” Again, whether or not this transpires may not be as important as students feeling that way in the first place.

However, aside from casual observation, is there an actual measurable impact of giving personalized feedback? It might be natural to expect students who perform poorly on a first exam to study more or seek additional help, so perhaps students already have enough motivation without feedback to work harder. Students who do well on exams may not feel like they need to change anything, and thus the impact of feedback on them may also be questionable.

To explore the answer to this question, I use data from a large state institution for a number of Principles of Microeconomics courses. Courses were taught by different professors over a 3-year period. Students in each class are placed into deciles based on their performance on the first exam, and again based on their performance of the second exam. Then, I find the difference between each student’s exam 1 and exam 2 deciles. A positive increase implies that the student improved on their performance relative to their classmates between exams. I find that classes where personalized feedback was given have students in the bottom 10th and 20th percentile and top 90th and 100th percentile perform better than peers in classes that do not give personalized feedback. Feedback does not seem to influence the middle 60 percent of students, as their change in relative performance is similar to those in classes without personalized feedback.

The rest of the paper is arranged as follows. Section 2 discusses the data used for this study. Section 3 describe the method that was used to give personalized feedback. Section 4 analyzes the performance of students from the first to second exam. Section 5 discusses potential issues with the analysis, and Section 6 concludes the paper.

2 Data

The data used currently includes 29 classes of Principles of Microeconomics taught by 3 different professors at a large state institution over a span of 3 years (Fall 2016 to Spring 2019). In sum, there were 6968 students in these 29 classes. I expect to have closer to 10,000 students by the time I am done with collecting data from more classes. Students were typically sophomores at the time they took Principles of Microeconomics. While there is no pre-requisite to take the course, many students do enter the course

having taken Principles of Macroeconomics because the course number for Principles of Macroeconomics is listed before the one for Principles of Microeconomics. The majority of students in these classes are Business, Liberal Arts, or Engineering majors.

Of the 29 classes in the sample, 16 utilized personalized feedback. These 16 classes accounted for 3949 students in the sample, and were all taught by the same professor. The other 13 classes did not have any form of official feedback mechanism for students, and were taught by two other professors. Of course, students of these two instructors can still go to office hours to receive feedback, but none was given on a large scale.

In the 16 classes where students received personalized feedback, students received two personalized emails from the professor over the course of the semester. The first email was sent about a week after the first exam. The email outlines first how they did on the exam from the professor's point of view, then goes on to discuss their homework and quiz scores in the class. Next, the email outlines how grades are calculated in the class and estimates for the student their current grade based on everything they have done so far. This is also where the professor gives encouragement to the students who are not doing well and outlines ways to bounce back, or applauds students who did well but encourages them not to slack. Finally, the email ends with an invitation to talk more about their current grades and exam score in office hours. An example of what such a feedback email looks like is given in the next section of this paper. The second feedback email focuses more on what students need to achieve on the final exam to get a specific grade in the class.

I will only focus on performance of the first and second exams for all classes in the sample. This is done for a few reasons. First, classes differ in how many exams were given. At minimum, classes had two exams and a final. Some classes had four exams and a final. Second, some classes had a policy that the lowest exam score can be dropped. This tends to be the case for classes with at least 3 exams and a final. I will assume that students tend not to use their drop on a first or second exam. For the most part, students do want to start a semester off well and not plan on dropping their first or second exams. Of course, it happens that students have difficult circumstances arise and decide to use a drop early, but I do believe that this happens infrequently and will not skew my analysis too much. In a more careful analysis of the data in a later version of this paper, I will analyze which exam was dropped in classes that offered a drop-the-lowest-exam policy and provide evidence for my assumption.

Students who did not take the first or second exam are excluded from the analysis. Typically, they are the students with a 0 score on either exam 1 or 2. This ensures that only students who have taken both the first and second exams in their class are analyzed. The averages on the first and second exam varies by class, but are always within 5 percentage points of each other for each class. There are variations in average exam scores by professors. Most have averages around 70%, while some have averages in the 50%*s*. Exam scores could be curved or adjusted. I discuss in the analysis section a way to hopefully get around the differences in how each student was graded.

3 Method

In classes that utilized personalized feedback, every student received an email after the first and second exams (these classes have two exams and a final). The first feedback

email focused on how a student was doing and how they could improve. The second feedback email discussed the student's performance in the class and gave estimates on what they needed receive on the final exam to achieve certain grades. Again, this paper will focus only on how the first feedback email affects student performance in a subsequent exam. Below is an example of a personalized feedback email that was sent to a student who scored poorly on the first exam. I replaced the student's name with my own for obvious privacy reasons.

Hello Kelvin – hope this email finds you well! Now that the first exam grade is posted, I'm emailing you to give you some feedback on how you are doing. Regarding the exam, your exam score was below the class average of around 70%. It looks like you expected to get a grade in the B range, based on your answers on the MobLab survey. It's definitely good to aim high, and I think that is still something that is doable for the second exam. However, something will have to change for that to happen. If you are up for it, I would love the opportunity to work with you more during office hours so we can ensure you do better on the second exam. I know you can achieve more and I want to make sure you know you have resources available to you for that. You had also mentioned on your MobLab survey that you were not able to do all the practice questions. Before the second exam, I would highly recommend doing all the practice questions since it will give you good practice for questions from the exam. And if anything is unclear, definitely come talk with me during office hours. As for your homework grade, it is looking great... make sure to keep that up. Regarding your quiz grade, it is also looking great. Make sure you also keep that up!

As I said in class, your current grade can vary depending on how you end up doing on your second exam and final. So, in this email, I will calculate your current grade both according to the syllabus (18% exam 1, 18% exam 2, 24% final, 20% homework, 20% quizzes) and also according to the grade reweigh I mentioned (10% exam 1, 22% exam 2, 28% final, 20% homework, 20% quizzes). But since we have only just completed week 8, I will take your grade proportional to the amount of work that we have done so far. So, this means for your grade estimate I use 10% for quizzes (for 8 weeks of quizzes), 7.27% for homework (4 homework out of 11, and dropping lowest two weeks' worth), and either 18% for exam 1 (syllabus grading scheme) or 10% for exam 1 (reweighted grading scheme). This will be out of either 35.27% or 27.27% of the current course grade possible, depending on which weighing scale was used.

If your grade is calculated according to the syllabus, it currently is 77.02, which is a C+. If it is calculated with exam 1 counted as 10%, then your current grade is 82.1, which is a B. This estimate already includes dropping your lowest two homework scores so far and any EconPoints you may have earned so far (i.e. does not include any discussion board post points, as those will only be tallied at the end of the semester). I will run your grades under both schemes at the end of the semester and take the higher one. Your overall grade at this point is higher than

what you might have expected (given your exam 1 score) because homework and quiz scores also play a huge role in your grades. Once again, I can't stress enough how important these things are at the end of the semester. With the way I reweight exams, there is definitely a great shot that you will still end up with a good grade in the class. However, you will want to have a plan on how you can do that as soon as possible. I would be very happy to talk with you more about this (and highly encourage you to!), so don't hesitate to schedule a time to meet with me. You can also ask me anything about material that is unclear to you, even if you think there is a lot that is unclear right now. Everyone has to start somewhere!

Please do not hesitate to ask any questions via email or during office hours. If you would like to further discuss your grade or anything in the class, please let me know and I would be happy to meet with you. Remember that you can sign up for office hour appointments if my regularly scheduled office hours do not work for you. To do so, just go to "Office hours" at the top of our Canvas course page for this class, and look through the guide to sign up for appointment slots.

Hope you are having a great Fall break, and hope to see you in class Wednesday. Rooting for you to do even better than you think you could!

To be clear, a lot of the things mentioned in the email were also discussed during class time. However, as students tend to miss a lot of what is said or simply forget, having it written on an email seems to be helpful. In the specific case of the classes using personalized feedback in this study, exams can be weighed differently to provide incentive for students to do better on future exams. Thus, there might be a higher chance that students are confused about how they are doing even though they are told how the reweigh would work.

To have a greater level of detail in these emails, the professor used surveys in class, done through an economics games platform called MobLab, to get data on students regarding how they are feeling about the exam and how they might improve upon their study habits. This survey was given right after the exam was taken, but before students know their exam scores. All of this was done to help students receive specific advice that can help them perform better in the class. Also, having more details may have the added impact of students feeling that the feedback was meant for them. This increases the chance that students will read through the email and take the advice to heart.

The professor in this study constructed these feedback emails using a mixture of data from surveys (as discussed previously), STATA, and a mail merge add-on in an e-mail client called Mozilla Thunderbird. The mail merge tool in Thunderbird makes sending personalized feedback emails less time consuming (and perhaps feasible). The mail merge add-on used is different from traditional mail merge tools in that numerous variables can be used. Thus, comments on how to improve in the class can be easily varied across students. For example, if a student answered that they did not utilize practice exams in the class, it can be programmed in STATA that an advice for this student is to be sure to take a look at the practice exams for the next exam. Students can also be separated by their performance on exams and other assessments (like

homework, quizzes, or participation) to give more specific advice. For instance, for students who did well on exams but not on homework, the feedback email gave them a nudge into taking the time to complete them.

Since the mail merge add-on takes a CSV file to send emails, the processing of data can also be done in Excel with nested IF functions or VLOOKUP instead of in STATA. I am sure that other options also exist for sending emails with many variables, or for compiling the data from students, but I only mention what was used for the feedback emails sent in this study.

4 Analysis

Since each professor curves and weighs exams differently, it is not ideal to look at actual grade percentages on exams. For instance, one class may have an average of 50% on the first exam and another class with an average of 70%. Yet, this may simply be a function of how the tests were curved. Whatever way that exams are graded, a similar grade distribution still is likely to occur at the end of the semester across classes taught by different professors.

Instead of looking at percentages on exams, I break each class, taught by each professor in each semester, into deciles based on their performances on the first exam. I also put students into deciles for the second exam. Then, for each student, I compute the difference in their decile between exam 1 to exam 2. For example, if a student was in the 7th decile (top 70% of the class) for the first exam and improved to the 8th decile for the second exam, she would have a decile difference of 1. A student who placed into a worse decile for the second exam than the first will have a negative difference in decile. A zero-decile difference is for students who placed in the same decile for the first and second exams.

I believe that this is a more effective way of comparing students from different classes because it factors in the difficulty of the exams in each class of each semester. Otherwise, a particular exam could have just been easier leading to most students having higher scores, or harder causing students to do worse in general. By using deciles, I can measure a student's performance on exams relative to the rest of their class.

Table 2 shows the result of a series of simple regressions where the difference in exam 1 and exam 2 deciles for each student is regressed on whether the class utilized personalized feedback by each exam 1 decile group. The constant for each regression represents the change in decile between exam 1 and exam 2 for students in the specific exam 1 decile group. The variable feedback is a dummy variable denoting whether the class utilized personalized feedback or not. For instance, the results show that students in the bottom 10% of the class on average will improve their performance by 1.21 deciles on the second exam. This likely points to the natural tendency for those who did poorly on an exam to work harder to do better. For this same group, in a class that feedback was given, students on average were 1.51 deciles higher on the second exam, which is .3 deciles better than students in a class without personalized feedback.

The result from this analysis suggests that those that are most helped by personalized feedback are those at the bottom (the 1st and 2nd decile) and those at the top (the 9th

and 10th decile). Even though the result is not statistically significant for the 2nd decile, it is just barely so. It appears like students in the middle 60% of the class are not as affected by personalized feedback.

Table 2: Difference in Exam 2 Deciles and Exam 1 Deciles, by Exam 1 Deciles

		Exam 1 Deciles, N=6027									
		10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
		(N=613)	(N=637)	(N=606)	(N=607)	(N=586)	(N=676)	(N=557)	(N=591)	(N=584)	(N=570)
feedback		.30** (.14)	.25 (.17)	.06 (.18)	-.17 (.20)	-.03 (.21)	.01 (.19)	.17 (.22)	-.06 (.19)	.31** (.19)	.26* (.15)
constant		1.21*** (.11)	1.01*** (.14)	.94*** (.14)	.56*** (.16)	.10 (.17)	-.44*** (.15)	-1*** (.18)	-1.09*** (.15)	-1.73*** (.15)	-1.46*** (.12)

Standard errors are in parenthesis.

*** denotes statistical significance at a 1% level, ** denotes significance at the 5% level, and * denotes significance at the 10% level.

Those who are at the bottom responding well to feedback may have been an expected result. These are students who may have studied inefficiently, or needed a nudge to work harder. The feedback therefore helps them understand what might have went wrong and points them to the right resources. A more surprising result may be those at the top deciles. Naturally, those who are in the top deciles have a difference that is negative because students cannot move above the 10th decile. The results do show that students in the top 2 deciles still benefit from feedback. This could point to the feedback emails being an encouragement for them to keep investing in the class, or perhaps students are heeding the warning to not let their first exam score decrease their effort in the class.

5 Discussion

There is one major issue in this analysis that may be obvious, and that is the lack of control for the professor who is teaching. Ideally, each professor teaches a mixture of classes that does not use and uses personalized feedback, but this is simply not the case in the data. Since only one professor in the data uses personalized feedback, the instructor effect simply cannot be disentangled from the analysis. A simple fix for this problem would be for the professors in the data to send feedback emails to random students, but this is currently not done.

Another issue is in the composition of students in each class. Students have information prior to the semester starting of who the professor is. Some students may prefer the teaching style of one professor over another, and some classes are filled up quicker than others because of this. It could be that students who are hard workers naturally sign up for a specific professor's class, whereas those who do not care as much sign up for whatever class works best with their schedule. In any case, this would lead to a dampened effect for the effectiveness of personalized feedback.

There are not controls for students other than their exam scores. While data on homework, quizzes, and participation is in the data for each student, I did not include them because I believe they all point to a similar thing. For instance, a student who was positively affected by personalized feedback will not only work harder on the next exam, but spend more time on homework, quizzes, and increase participation. What I do plan on doing in a future version of this paper is to go through each class and track homework, quiz, and participation scores prior to exam 1 and after exam 1 to see

whether personalized feedback also affects performance on those assessments. Lastly, I do not have access to student academic records to add additional controls.

6 Conclusion

In this paper, I discuss the importance of giving students feedback over the course of a semester and outline a way to do so in large classes. The analysis shows that students at the bottom and top 20% after the first exam seem to be most helped by personalized feedback. Even though the analysis suffers from an identification issue and the data from a composition problem, the initial results are encouraging because it shows that some students really are positively impacted. At the end of the day, many students do believe that this feedback helped them and that they were more invested in the class because of them. In my opinion, that in itself is reason enough to justify the costs.