

Measuring the Effect of Aplia on Cadet Performance in the Principles of Economics

Course

Majors Mike Shekleton and Bill Skimmyhorn

Center for Teaching Excellence, Master Teacher Program

4 April 2008

Abstract

This paper looks at the impact of using an online homework software solution, Aplia, on cadet performance during in-class graded events and on instructor satisfaction with teaching SS201. While we don't have a causal argument, we find that an increase in Aplia homework performance by 1 percentage point results in an increase in in-class graded performance on the multiple choice portion of an exam from 0.161 percentage points (Microeconomics WPR) to 0.211 percentage points (Macroeconomics WPR). For the non-multiple choice portion of an exam, an increase in Aplia homework performance by 1 percentage point results in an increase in in-class performance from 0.036 percentage points (TEE) to 0.105 percentage points (Macroeconomics WPR). This indicates that the increase in in-class graded exam performance may simply be a test-taking effect, since the practical significance of the impact on multiple choice performance is better. Future research that is on-going will utilize a treatment and a control group to allow the authors to ascertain a causal link between use of an online homework software solution and in-class graded examination performance. Lastly, while this paper focuses on the impact on cadet performance, instructors should also consider how the use of an online homework software solution frees up their resources, mainly time. To assess these effects and to gain greater insight into how best to use technology in the economics classroom, surveys were administered to all of the instructors teaching SS201 in AY08-02. Our future work will also include surveys of all cadets enrolled in SS201 in order to determine their receptiveness to Aplia as a teaching tool and to solicit their feedback on better incorporation of technology as part of the teaching of economics.

Background

Because of the specific nature of the academic environment at USMA and that of the core economics course, it is worthwhile to spend some time describing this environment as it affects some of our thoughts *a priori*. Cadets typically take SS201, Principles of Economics, (or SS251, Advanced Principles of Economics) during their second year at the Academy, with the only co-requisite requirement for the course being MA104, Univariate Calculus. Classes are held every other weekday with 55 minute class periods, and the class size is small, with generally 14-18 cadets in a class. There are 40 lessons during the course of the semester.

Cadets are required to complete an Aplia assignment for every lesson where new material is presented, and the questions that are used are from the Mankiw *Principles of Economics* question bank designed and provided by Aplia. Redundant questions from the assignment are deleted to ensure that cadets can complete both the reading assignment and the Aplia assignment within the two hours of time allotted for each lesson outside of class. Assignments typically range between 10-15 questions (with each part of multipart questions counting as an individual question in this number). USMA uses the “Thayer System” where cadets are expected to learn the material prior to class and then come to class with questions about what they didn’t understand. Thus, the Aplia assignments are completed prior to instructors presenting the material to cadets.

Data

Our primary dataset includes 2,117 potential observations across four semesters during the 2006 and 2007 academic years. This data is drawn from Aplia (both assignment scores and scores from the multiple choice portion of exams) as well as from the Academy Management System. Our dependent variable in all models is a cadet’s performance on an in-class graded exam. We look specifically at the performance on the multiple choice part in one family of models, and another family of models looks at the performance on the non-multiple choice part of the exams.

For the multiple choice portion of exam scores, we have data only for those observations where this was conducted via Aplia in the classroom during the test. However, not all cadets were able to complete their multiple choice portion of exams online via Aplia due to hardware issues or network issues. Our argument is that these cadets are a random selection of cadets and does not bias our dataset towards any particular outcome. Another limitation in the data is with the Macroeconomic exam, where only one semester's worth of data exists due to the absence of the course director that implemented Aplia into the course, and the decision by the following course director to suspend online testing due to the perception that cadet's performances were hindered versus those that were able to take the multiple choice portion on a hardcopy test.

We will now look at the specific variables in the data set that will be used in our models:

Micro Aplia % - This is a continuous variable that measures a cadet's performance on Aplia homework during the microeconomics block.

Macro Aplia % - This is a continuous variable that measures a cadet's performance on Aplia homework during the macroeconomics block.

SS201 Aplia % - This is a continuous variable that measures a cadet's performance on Aplia homework during the entire semester (micro, macro, and international blocks).

Credit-hours – This is a continuous variable that measures the course load of a cadet during the semester that he/she took SS201/251.

WPR 1 MC % - This is a continuous variable that measures the % a cadet scored on the multiple choice portion of WPR 1.

WPR 2 MC % - This is a continuous variable that measures the % a cadet scored on the multiple choice portion of WPR 2.

TEE MC % - This is a continuous variable that measures the % a cadet scored on the multiple choice portion of the TEE.

WPR 1 Not_MC % - This is a continuous variable that measures the % a cadet scored on the non-multiple choice portion of WPR 1.

WPR 2 Not_MC % - This is a continuous variable that measures the % a cadet scored on the non-multiple choice portion of WPR 2.

TEE Not_MC % - This is a continuous variable that measures the % a cadet scored on the non-multiple choice portion of TEE.

MA100 – Cadets who have a weaker mathematics background are enrolled in a pre-calculus course prior to the standard four course, core math sequence. Thus, we expect that cadets who have taken MA100 will have more difficulty in the course, both during the firm theory block where we integrate differential calculus into profit maximization problems and simply with the use of graphical model, and so we will control for this using this dummy variable.

PrevCQPA – This is a continuous variable that measures the cadet's Cumulative Quality Points Average at the beginning of the semester. The CQPA is a combination of their Academic, Military and Physical performance scores. This variable captures how well a cadet performs in the Academy's demanding environment.¹

Instructor – This is a dummy variable to control for variances across instructors.

AY – This is a dummy variable to control for variances across academic years.

AT – This is a dummy variable to control for variances across academic terms.

Afternoon – This is a dummy variable that captures whether a cadet had class during the morning or early afternoon. We use this variable to control for performance variation due to the time of day when a cadet was in class.

Black – This is a dummy variable to control for variances due to race/ethnicity

Hispanic - This is a dummy variable to control for variances due to race/ethnicity

Other - This is a dummy variable to control for variances due to race/ethnicity

Sex – This is a dummy variable to control for variances due to gender

¹ $CQPA = 0.65 \times APS + 0.25 \times MPS + 0.10 \times PPS$

Active – This is a dummy variable to control for an instructor’s support of the integration of Aplia into the classroom

<u>Descriptive Statistics</u>		
Variable	Mean	Std dev
PrevCQPA	2.896979	0.5518017
Credit Hours	18.94048	1.336294
Micro Aplia %	72.89683	15.41422
Macro Aplia %	69.72767	16.95662
SS201 Aplia %	67.94736	15.31628
WPR 1 MC %	72.08347	15.5313
WPR 2 MC %	67.59554	16.16043
TEE MC %	76.52797	11.78003
WPR 1 Not_MC %	85.19476	10.8481
WPR 2 Not_MC %	85.8936	10.11665
TEE Not_MC %	87.57834	8.805898

In addition to the data on cadet performance, we have designed a survey that will be completed by all cadets in SS201 & SS251 during AY08-2. This survey aims to measure cadet perceptions of the effectiveness of Aplia, possible alternate uses of Aplia, and other student feedback on the use of Aplia software in the course. This feedback provides an excellent complement to the quantitative data gathered above. It may be the case the Aplia (or other) software use is justified, even if not on the basis of cadet performance. Cadets may perform equally well with or without the software, but they may prefer (or not prefer) to use Aplia. We also expect to gain greater insight into how best to use the software (ie, before class as a

instructors' assessments of the effectiveness of Aplia, as well as allow them to suggest alternate uses of Aplia (pre-class, post-class, or in class) and/or alternate software and/or homework assignments. This data is important in making a decision on whether or not to use Aplia, as instructor preferences and perceptions are important considerations when designing a course.

The survey is included below:

**SS201 Instructor Feedback
Aplia Survey**

Please take a few minutes to answer these questions and share your comments on the use of Aplia in SS201.

1. Would you recommend using Aplia as a teaching tool in economics? Y N

2. Would you choose to use Aplia if you were the only instructor? Y N

3. Did you require students to complete Aplia in SS201? Y N Some

4. How many points did you allocate to Aplia in SS201? _____ / 1000 points

5. Do you think requiring Aplia completion on a daily basis trades off with:

Reading? Y N

Other studying techniques? Y N

6. Would you prefer to use Aplia as a method of teaching students new material (pre-class) or as an application method (post-class)? _____

7. What did you enjoy most about using Aplia? (list as many things as you can think of)

8. What did you enjoy least about using Aplia? (list as many things as you think of)

9. Would you prefer to use homework assignments from? (circle) Aplia Textbook Other (explain)

10. Do you recommend using the experiments in Aplia during class? Y N N/A
(If Yes, Please list the experiments you find the most useful)

11. How could Aplia be better utilized in SS201?

12. Do you have any other comments on Aplia use in SS201? _____

The Models

We employ three families of models (which each “family” looking separately at WPR 1, WPR 2, and the TEE). The first family of models looks at how cadet performance on the multiple choice portion of an exam varies with Aplia performance. The second family of models looks at how cadet performance on the non-multiple choice portion of an exam varies with Aplia performance. The third family of models uses an instrumental variable (IV) to strip out potential endogeneity issues from how instructors’ support for the Aplia software may impact a cadet’s performance on Aplia, and hence, Aplia’s impact on cadet graded performance.

Our first family of models looks strictly at how performance on the graded Aplia problem sets correlated with performance on the multiple choice portion of the in-class graded exams. Thus, we look at the correlation between the performance on the microeconomic problems sets with the performance on the multiple choice portion of the microeconomic WPR (WPR1), between the performance on the macroeconomic problem sets with the performance on the macroeconomic WPR (WPR2), and between the performance on all problems sets and the term-end examination (TEE, which is a cumulative examination).

We used a level-level specification to determine the impact of increased exam performance based on our variable of interest, the performance on the corresponding Aplia assignments.

$$\begin{aligned} WPR1\ MC\% = & \beta_0 + \beta_1\ Aplia\ Micro\ \% + \beta_2\ Credit\text{-}hours + \beta_3\ PrevCQPA + \delta_{1-13}\ Instructor \\ & + \delta_{14}\ Rock\ Math + \delta_{15}\ Class\ Period + \delta_{16-18}\ Minority + \delta_{19}\ Sex + \\ & \delta_{20-21}\ AYT + u \end{aligned}$$

$$WPR2\ MC\% = Aplia\ Macro\ \% \qquad TEE\ MC\% = Aplia\ SS201\ \%$$

Our second family of models used the same controls and procedures as above, except that instead of looking at Aplia’s impact on multiple choice performance, we instead looked at the impact on the non-multiple choice portion of these exams. The hypothesis that we were testing here was to see if Aplia provided a learning effect (in which case, we should see a

similar change in exam performance) or a test-taking effect (in which case, we should see a greater increase in multiple choice grades than in non-multiple choice grades). In other words, using a multiple choice based homework system could simply provide an edge in taking multiple choice portions of examinations.

Our third family of models that we looked at was an attempt to alleviate the endogeneity of how instructors' support for Aplia might bias cadet performance on Aplia homework assignments, and thus bias the true impact that Aplia might have on in-class graded exams.

First-stage

$$\begin{aligned} \text{Aplia Micro \%} = & \pi_0 + \pi_1 \text{PrevCQPA} + \pi_2 \text{Credit-hours} + \pi_3 \text{Afternoon} + \pi_4 \text{AT} + \pi_5 \text{Sex} + \pi_6 \text{Black} \\ & + \pi_7 \text{Hispanic} + \pi_8 \text{Other} + \pi_9 \text{MA100} + \pi_{10} \text{active} + v \end{aligned}$$

Second-stage

$$\begin{aligned} \text{WPR1 MC \%} = & \beta_0 + \beta_1 \overline{\text{ApliaMicro\%}} + \beta_2 \text{PrevGPA} + \beta_3 \text{Credit-hours} + \delta_{1-13} \text{Instructor} + \\ & \delta_{14} \text{Rock Math} + \delta_{15} \text{Class Period} + \delta_{16-18} \text{Minority} + \delta_{19} \text{Sex} + \delta_{20-21} \text{AYT} + u \end{aligned}$$

Results

Our models that there is correlation between a cadet's performance on Aplia homework and in-class graded exams. For every percentage point increase in a cadet's Aplia performance in the microeconomics block corresponds to a 0.16 percentage point increase in their performance on the multiple choice portion of the microeconomics WPR. For the macroeconomics block, a one percentage point increase in performance on Aplia corresponds to a 0.21 percentage point increase on the multiple choice portion of the macroeconomics WPR. For the course overall, a one percentage point increase in Aplia performance across the semester results in a 0.18 percentage point increase on the multiple choice portion of the TEE.

When we use similar models to test the correlation between Aplia performance and the non-multiple-choice portion of the in-class graded exam, we find a positive correlation; however, the practical significance is much less. A percentage point increase on Aplia during the

microeconomics blocks correlates to only a 0.04 percentage point increase on their non-multiple choice portion of the WPR. A percentage point increase on Aplia during the macroeconomics block correlates to a 0.10 percentage point increase on their non-multiple choice portion of the WPR. A percentage point increase on Aplia across the semester correlates to only a 0.3 percentage point increase on the non-multiple choice portion of the TEE.

While we cannot test whether the Aplia performance coefficients for the multiple choice and non-multiple choice portions of each exam are different in terms of statistical significance, the relative magnitude of the difference lends some support to a conclusion that Aplia provides more of a test-taking effect than a learning effect. In other words, as a multiple choice/fill-in-the-blank homework format, Aplia potentially provides an advantage when taking the multiple choice portion of an exam, an advantage that diminishes when cadets must recall information and provide explanations.

Our final family of models employs an instrumental variable to attempt to alleviate potential endogeneity. Therefore, we instrument *active*, which is a dummy variable that describes whether an instructor was an active supporter of using the Aplia software as the homework solution for the course, for *Aplia Micro %*. Using Two-Stage Least Squares (2SLS), we run the first-stage regression on the reduced form equation and then test the exclusion restriction, that $\pi_{10} \neq 0$. The P-value for t-stat is 0.135, which means that we fail to reject the null hypothesis that $\pi_{10} = 0$. *Active* and *Aplia %* are not correlated, which means that we cannot move on to the second stage, where we would run the regression using the *Aplia Micro%* estimator that has been stripped of the endogenous effect of whether an instructor was an active Aplia supporter or not.

Because of the lack of observations for WPR 2 multiple choice scores during the 2006 academic year, we cannot test a model for WPR 2.

We completed the same process for the TEE, using the instrument *active* for the endogenous variable, *SS201 Aplia %*. During the first-stage regression, we find a P-value for

the t-stat to be 0.002, which means that we cannot reject the null hypothesis that *active* and *SS201 Aplia %* are not correlated; we can continue to the second-stage regression. Here we find that our variable of interest is not significant, an indication that we have a weak instrument (loosely defined as the problem of “low” (not zero) correlation between the instrument and instrumented variable, which are *active* and *SS201 Aplia %*). In this case the coefficient on *active* is 0.02.

Survey Results

The results from our qualitative surveys are easy to interpret. The table below displays the results of the surveys issued to the instructors for AY08-02 and demonstrates a significant but not majority level of support for Aplia use and some common recommendations for improvements in the use of Aplia in SS201. This survey was issued to 12 instructors and 11 completed the survey, resulting in a 92% response rate.

Question	Yes	No	
Would you recommend using Aplia as a teaching tool?	73%	27%	
	Yes	No	
Would you choose Aplia if you were the only instructor?	45%	54%	
	Yes	No	Some
Did you require your students to complete Aplia problems?	45%	9%	45%
	Yes	No	
Do you think Aplia homework trades-off with assigned reading?	73%	27%	
	Yes	No	
Do you think Aplia homework trades-off with other SS201 studying?	45%	45%	
	Yes	No	N/A
Do you recommend using the Aplia experiments during class?	36%	45%	18%
	Pre-Class	Post-Class	
When would you prefer to use Aplia pre-class or post-class?	18%	82%	
	Aplia	Text	Other
Would you prefer to assign homework from Aplia, the textbook, or something else?	55%	18%	27%

Note: Percentages may not sum to 100% due to rounding.

We also solicited specific feedback from instructors on their thoughts on Aplia and obtained the following results:

What did you enjoy most about using Aplia?

Response	Qty
Automatically graded	6
Assessing student challenges	4
Incentive to complete homework	4
Practice problems offer feedback	2
Existing bank of questions	1
Experiments	1
Tracking participation	1
Opportunity to work with graphs	1

What did you enjoy least about using Aplia?

Response	Qty
Poorly written questions	6
Site connectivity	3
Delayed feedback	2
Organizing times on the website	2
Different questions than our WPR questions	1
Too frequent	1
Non-intuitive format online	1

How could Aplia be better employed in SS201?

Response	Qty
Post-class assignments	5
More dynamic feedback	2
More infrequent assignments	1
In conjunction with book questions	1
Drop experiments	1
More experiments	1
Pare down daily problem sets	1
Work a few Aplia problems in class	1
Develop board/desk problems from Aplia problems	1
Assign pre-class but grade post-class	1

Caveats and Future Work

Because of some of the unique characteristics of the learning model at West Point (small class size, completion of assignments prior to instructor presentation of the material, and time constraints on cadet schedules), these results may not apply well in other settings. Additionally, there are some specific characteristics of Aplia during the study period that may affect the value added of using an online-homework platform. For example, during the study period, assignment answers could be changed at any time by the cadet prior to the assignment due date and time. Once that time came, the final recorded answer was graded. While this feature provides maximum flexibility in completing an assignment, if a cadet didn't go back after the due date and time to check individual assignment questions, then they don't receive the necessary feedback to learn from their mistakes. In contrast, a system that grades a student's first response and then provides immediate feedback ensures that a student at least has the opportunity to learn from any incorrect responses.

While this paper used data mining of prior semesters to attempt to establish correlation, the SS201 course is currently using a natural experiment that takes advantage of the random assignment of cadets to their class section to get at a causal link. We solicited volunteers from the pool of SS201 instructors to participate in this experiment. We assigned each of the volunteers (7), to teach half of their sections (1 out of 2 sections or 2 out of 4) with zero points assigned to daily Aplia homework (0% of the final grade) and to teach the other half of their sections with 100 points assigned to daily Aplia homework (10% of the final grade). We intend to evaluate these different point allocations on cadet's performance on the examinations in SS201.

In addition to our evaluation of the effects of point assignment to Aplia on the major graded events in SS201, we will also evaluate the effects of point assignments on cadet learning as measured by the Test of Understanding College Economics (TUCE). We have administered the Micro and Macro pretests and will administer the posttests after completion of

the Micro and Macro blocks of instruction. This model provides for the clearest interpretation of causality as we can more confidently control for Aplia completion and instructor variation (i.e., instructors that don't assign Aplia will use other assignments and methods that would be used in the absence of Aplia) and their corresponding effects on cadet performance. We will complete this research during the summer of 2008.

Conclusion

Our models demonstrate that Aplia performance correlates positively with in-class graded exam performance, although this correlation isn't strong, and in some cases is quite weak. Without a causal argument, we are unable to state a definitive conclusion whether Aplia truly improves performance or is simply correlated with some other omitted variable that is the true causal mechanism (e.g., maybe instructors can prepare better examples or board problems for class since they don't have to commit any time to grading homework). However, if the natural experiment that is being conducted this semester turns up similar results, then even though the effect might be weak, because of the time that the online homework platform saves instructors (i.e., problems sets/homework assignments that don't need to be created and/or hand graded is a significant factor), the use of such a platform may provide more benefits than if you were to solely consider its impact on cadet performance.

Our survey results provide important data and complement our quantitative analysis in this regard. Our results indicate that most instructors do recommend using Aplia and that most did require students to complete Aplia homework. However, most instructors would prefer to use Aplia as an application tool (ie, post-class) instead of as a pre-class homework assignment. In addition, we gained some excellent recommendations on how to better utilize Aplia in SS201. This feedback will allow us to improve the design and administration of SS201, to enhance cadet learning, and to improve instructor satisfaction in the course.

Model 1

	Micro WPR MC %	Macro WPR MC %	TEE MC %
Micro Aplia %	16.11784*** (3.635274)		
Macro Aplia %		21.11574*** (5.26)	
SS201 Aplia %			17.98919*** (2.438941)
Term Credit Hours	-1.095099*** (0.2969561)	-1.177104* (0.6792398)	
Previous CQPA	11.83012*** (0.9973566)	11.31674*** (0.6792398)	9.539087*** (0.7553414)
MA100 (1=yes)	-4.821046* (2.474289)	-9.454177*** (3.59694)	-5.58269*** (1.586991)
Term (1=fall semester)	1.69622** (0.8544599)		
Sex (1=female)			-2.672087*** (0.7992449)
Other (1=yes)			-2.139458*** (0.6463814)
Constant	46.02916 (5.872875)***	44.36836*** (14.10624)	38.58823*** (2.254839)
R-squared	0.3086	0.3454	0.3862
# observation	1006	311	1141

***significant 1%, **significant 5%, *significant 10%

Model 2

	Micro WPR non-MC %	Macro WPR non-MC %	TEE non-MC %
Micro Aplia %	4.355005* (2.551828)		
Macro Aplia %		10.16797*** (3.843461)	
TEE Aplia %			3.462108* (1.771389)
Term Credit Hours	-0.4025085** (0.1922893)	-0.2702387 (0.3872562)	-0.3776501** (0.1780088)
Previous CQPA	10.15722*** (0.792744)	8.393256*** (1.06271)	8.41515*** (0.473163)
MA100	-4.284392** (1.771831)		-3.434068*** (1.019937)
Afternoon	-1.289259** (0.5781253)		-1.09639** (0.4295164)
AT	-1.830613*** (0.6686847)		-1.227927** (0.5387774)
Sex	-2.285354*** (0.7540335)		-2.342502*** (0.5019832)
Other	-1.497967** (0.6828828)		
Constant	64.52532*** (3.744856)	63.36487*** (8.376421)	72.41808*** (3.601251)
R-squared	0.3717	0.3683	0.4244
# observation	1041	366	1149

***significant 1%, **significant 5%, *significant 10%