Does Calculator Use and Test Format Mask Weakness in Basic Math Ability? Experimental Evidence in Principles of Economics

By Melanie Allwine and Irene Foster*

Results from an experiment in Fall 2013 of 971 incoming students at George Washington University are reported. In this experiment after students were given an assessment to ensure they had the necessary Algebra I skills to take a Principles of Economics course, they were randomly allocated to a treatment or control group to test if there was a significant impact of test format, calculator use and type, and the interaction of calculator use/ type and test format on students scores. The results from this experiment do suggest that each treatment had a significant impact on students scores, with much variation depending on the type of question asked.

In results found in Allwine and Foster (2014) it was shown that the Algebra I Assessment was a good predictor of how well a student did in Principles of Economics. It was predicted that students that passed on their first attempt scored 12 percentage points higher than students that failed (significant at the 10% level). However, the SAT Math was found to have little explanatory power with respect to students performance in Principles of Economics, despite the fact that the SAT Math and Algebra I Assessment covered the same material. Allwine and Foster (2014) hypothesized that the reason for this difference in explanatory power was due to the difference in test format and/or calculator use on the assessment and SAT Math.

In this paper we test this hypothesis by running a randomized experiment in George Washington Universitys Principles of Microeconomics course, randomly assigning students to use and of one of 3 types of calculator (basic, graphing , own calculator) and then randomly assigning students a test format (multiple choice or open-ended) in order to analyze not only the impact of calculator use/type and test format on students mathematical skills necessary to do well in a Principles of Economics course displayed through an algebra i assessment but also the impact of the interaction of test format and calculator type/use.

In Principles of Economics, Algebra I is considered to be a prerequisite at George Washington University. The Algebra I Assessment is given to all incoming Principles of Economics students to test their Algebra I skills, as it has been

* Corresponding author Foster: George Washington University, 2115 G St. NW Suite 340, Washington, DC, 20052, fosterir@gwu.edu. Allwine: George Washington University, 2115 G St. NW Suite 340, Washington, DC, 20052, melanieallwine@gmail.com.

found by professors that students without these skills perform poorly in the class. This was then proven to be the case in Allwine and Foster (2014). The assessment is made up of 20 questions that had a direct relationship to mathematical understanding in economics. The questions cover the same material as the SAT Math. Results upon first giving the assessment were astounding, with a failure rate of over 50%.

The layout of the experiment was to give students an additional assessment (referred to as a questionnaire) after the Algebra I Assessment to test if indeed our hypothesis that test format, calculator use and type, and/or the interaction of the two were impacting students mathematical abilities as demonstrated through their score on the assessment. The experimental design was to randomly assign students (using their respective teaching assistant, as each assessment was administered by teaching assistants during their respective discussion section) to the use/ type of calculator, conditional on students SAT math scores as a proxy of students innate math ability coming into the course. Randomization is done in this way to ensure that there was compliance of students with the restrictions of calculator use/ type and to reduce any unneeded stress on students. To further test any interactions between calculator use and test format, individual students are randomly assigned to a specific test framework (again conditional on students' SAT math scores).

The empirical strategy is to use a difference in difference framework to discern the impact hat the test framework and calculator use have on total score of students on the questionnaire, as well as the impact on whether students answered individual questions correctly from treatment. We then also use a difference in difference strategy to test for any interactive impacts of calculator use and test format.

Results suggest that calculator use/ type and test format have an impact on students overall scores on the questionnaire, with an interactive impact from adding a graphing calculator to the multiple choice framework. These results on the overall score are drive by the impact of calculator use/type and test format on students' abilities to answer individual types of questions. We find that for questions relating to Algebra and Functions where a word problem is used and students are expected to write out the functional form and solve for the two unknowns, calculator use and multiple choice alone have similar impacts, while adding a graphing calculator to the multiple choice framework has a large impact on the percentage of students that correctly answer the problem. Results also suggest an increasing return to test format and calculator use/ type. Whereas in the case of the percentage change and data analysis problems, multiple choice had the largest impact, likely because students that have a sense of numeracy can analyze the possible answers and determine which answers are most likely correct when given multiple potential answers.

The layout of the rest of this paper is as follows. Section 2 gives a literature review. Section 3 describes the Algebra I Assessment. Section 4 describes the layout of the experiment. Section 5 summarizes the assessment results for those students that partook in the experiment. Section 6 describes the empirical strategy. Section 7 presents results. Section 8 concludes.

1. Literature Review

The use of calculators in the learning of mathematics has been looked at for many years. The research is so vast that it will not all be discussed here. We will summarize some of the more recent results that have been found on the impact of calculators on SAT math questions and ignore the impact that calculators have on the acquisition of math skills, though we note that if calculator use does impact the acquisition of math skills that this could be causing a differential impact of calculator use on the ability of students to correctly answer SAT math questions.

Bridgeman, Harvey and Braswell (1995) analyze the impact that calculator use has on the ability to correctly answer SAT math questions. They find that the benefit/cost of using a calculator depends on the type of SAT math question asked. The effect of calculator use on a student's ability to correctly answer a question ranged from positive through neutral to negative and that these range of effects were present in both difficult and easy questions. This follows our own results below. **?** evaluate questions included in the SAT math in 1996 and 1997 and find that students with better math skills are more likely to use a calculator and therefore score better.

Using ones own calculator may significantly impact a student's ability to correctly answer a SAT math question. Calculators vary in functionality and being familiar with a calculator yields benefits. Bridgeman and Potenza (1998) compare scores on SAT math questions when students use their own calculator versus an on-screen calculator. They find no impact of using the provided versus students' own calculators.

The impact of test format on test scores is another area of research that is vast, so again we will focus of the imact of test format on the ability of students to correctly answer SAT math questions as well as the impact on students ability to answer economics questions correctly. The SAT math contains questions that are both open-ended and multiple choice. From our research we have not found any analysis as to the difference in students performance on multiple choice vs. open ended questions in the SAT math. This may be because in the SAT math the questions that are asked using a multiple choice format are meant to assess students on different skills than those that are asked in an open-ended format. There is also no information available on scores from the SAT math on multiple choice versus non-multiple choice questions so it is difficult to analyze any impact on the actual exam.

Becker and Johnston (1999) look at the ability of multiple-choice and open-ended questions to evaluate students economic knowledge. Since the questions chosen from the SAT math in our assessment are chosen based off the skills necessary to do well in a principles of economic course, this research is pertinent. The authors find that both types of questions show different dimensions of knowledge and so both could be used for testing in economics. **?** find similar results.

From our search, we have not found any work that has been done on the interaction of calculator use and test format on students ability to answer SAT math questions or economics questions.

From our search, we have not found any other university that is requiring students to be take a test to evaluate them on their mathematical skills before entering principles of economics. In mathematics and the physical sciences, it is common for students to be assessed on thei math skills before the beginning of a course. The Mathematics Diagnostic Exam is given to incoming mathematics students at New York Polytechnic. Bryan Mawr allows students to test their math readiness for a Physics course (though it is not required).

Many researchers have found that math skills are important to students performance in economics courses. Ballard and Johnson (2004) find that a student's acquisition of basic math skills is an important factor for performance in introductory microeconomics. Pozo and Stull (2006) run an experiment where the treatment group was given a grade incentive to complete a math skills unit. They find students in the treatment group performed better in the course than the control group. They also find a larger gain for students lower in the grade distribution. Lagerlof and Seltzer (2009) examine the impact of a remedial math course on students performance in economics, and find that though secondary mathematics has a great deal of explanatory power in determining how well a student does in economics, taking a remedial math course dud bit demonstrate a positive effect. Mallik and Lodewijks (2010) look at the correlation of variables with students performance in economics. They find that high school math scores (beyond general math) have a positive correlation with students performance in economics. Owen (2012) does a review of additional findings that mathematical skills impact students performance in economics.

2. Algebra I Assessment

In Principles of Economics, Algebra I is considered to be a prerequisite at George Washington University. Whereas in the past GW economics professors have assumed that students were proficient in Algebra I, faculty members observed that students were weak in mathematics and this inhibited their learning of economics. It was for this reason that in the Fall of 2010, principles faculty began giving an Algebra I Assessment to all incoming principles of economics students that counts towards 10% of each student's grade, demonstrating the importance of Algebra I in the understand of principles courses.[1]

Students are told (by email) of the assessment 2 weeks before classes start and are told that the duration of the assessment is 30 minutes, that no calculators are to be used on the assessment, and that the student must receive an 80% or

[1]Students receive 10% if they pass on any of the attempts and 0% if they fail all three attempts.

higher to pass (as this demonstrates proficiency in Algebra I). The assessment is administered in students discussion sections with a respective teaching assistant for the course. Students are given 3 chances total to pass the assessment. The department provides remedial math support for all students that fail the first assessment given.

The layout of the Algebra I Assessment is as follows: 20 questions are chosen from the SAT Math that had a direct relationship to mathematical understanding in economics. Each question as well as a summary of how the question relates to economics is listed below in Table 1. Questions were chosen that were simple and could be easily done without the use of a calculator. The questions chosen covered the same topics as the SAT Math within the same proportions, including: Numerics and Operations (20-25%), Algebra and Functions (35-40%), Geometry and Measurement (25-30%), and Data Analysis (10-15%). A copy of the assessment is available upon request.

A total of 1,354 students took the Algebra I Assessment in the Fall of 2013. Of these students the average total score was 13.5/20 implying that on average students answered 68% of the 20 questions asked correctly. To pass students were required to correctly answer 80% of the question correctly, implying that a 16/20 or above was a passing score. This led to a failure rate of 65% on the first attempt at the assessment.

3. Experiment

Immediately after the Algebra I Assessment was given, 902 students[2] from 2 professors of Principles of Microeconomics courses were given an Algebra I Questionnaire containing a sample of 10 questions from the Algebra I Assessment.[3] This sample of students included the students of 13 teaching assistants where each teaching assistant had 3 discussion sections with approximately 25 students in each, implying that 39 discussion sections were sampled. Of these 902 students, 7% did not consent for their information to be used in the experiment, leaving us with 831 students in our sample. Table 1 demonstrates that the share of students that did not consent does not significantly differ across treatment and control groups. This ensures that there is no bias in the results from students not wanting to participate based off of their assignment into a specific treatment or control.

Figure 1 explains the experimental design. First, each teaching assistant was randomly assigned to a calculator treatment or control conditional on students average SAT Math score. Randomization was first done at the teaching assistant level to ensure that treatment was followed as it would have caused students

[2] Of the 1144 students enrolled in the large principles of microeconomics courses 242 were dropped from the sample. These students were in a course with a professor who did not participate in the experiment. This could potentially bias our results in their application to all GWU students.

[3] The sample of questions maintained the same proportions as the Algebra I Assessment and SAT Math.

unneeded stress and would have been difficult to enforce if randomization was done by the discussion section or student. Each teaching assistant was also given a helper to ensure that each of the teaching assistant's classes would be strictly given the treatment or control. Randomization was conditional on the teaching assistant's students' average SAT Math score to ensure that treatment and control groups had similar math skill levels. Of the 39 discussion sections, 12 discussion sections (including 4 teaching assistants with 256 students) were assigned to the control group which did not have the use of a calculator, 9 discussion sections (3 teaching assistants) each were assigned calculator treatment in the form of being given a basic calculator at the beginning of the Algebra I Questionnaire, being given a TI-83 graphing calculator, or being allowed to use their own calculator.

The basic calculator used in this assessment were the same as those available to be used by the GW Economics faculty. Only addition, subtraction, multiplication, and division can be done with these calculators. There is no exponential function or square root. For students that were allocated into the treatment where they could use their own calculator, students were asked to report if the calculator was a basic calculator or graphing calculator. It was reported that 40% of students used a basic calculator, 56% used a graphing calculator, and 4% used no calculator. All students were advised previous to the assessment that they should bring a calculator, however it is likely that many students forgot the day of the assessment. It is for this reason that the results that are likely to be the most comparable from the Questionnaire to the SAT Math are those where students were given a graphing calculator as it is engrained in students that they should bring a graphing calculator to the SAT Math. It is also likely the most comparable because no checks were done for possible functions in students graphing calculators whereas the graphing calculators given to students had no functions downloaded onto them. The SAT Math checks all students calculators for downloaded functions.

To further test any interactions between calculator use and test format, we further randomize within each of the discussion sections (again conditional on students' SAT Math scores) such that half of students were given an open ended test framework and the other half of students were given a multiple choice framework. This was done by giving students an envelope with the assigned questionnaire inside.

The validity of the experimental design can be seen in Table 2, where statistically significant differences are tested between treatment and control groups. The only treatment group where students had a significantly lower SAT Math score in comparison to the control group was for those that received multiple choice and a graphing calculator (significant at the 10% level).

4. Assessment Results for Experiment Participants

Assessment results for the 831 students that participated in the experiment can be found in Table 4. Of the 831 students that participated in the experiment their average total score on the Algebra I Assessment was 13.4/20 implying that on average students answered 67% of the 20 questions asked correctly. This led to a failure rate of 66% on the first attempt at the assessment.

Table 4 also gives the average total score on the assessment for these students if only the 10 questions used in the Questionnaire had been used on the Algebra I Assessment instead of the full set of 20 questions. The average total score if only these 10 questions had been used would have been 7.7/10 implying that on average students answered 77% of these 10 questions correctly on the assessment. This would have resulted in a failure rate of 61.5%. The reason for the difference in average score and failure rate between those questions chosen for the questionnaire and those used in the assessment is that only 22% of student correctly answered question 9 on the assessment (not used on the questionnaire).

Table 4 lastly gives the percentage of students that correctly answered each of the questions on the assessment as well as the corresponding question number on the questionnaire. Questions 2 and 4 from the assessment were used on the questionnaire and covered Numerics and Operations. On question 2 of the assessment, 89.6% of students answered the question correctly and on question 4 of the assessment, 96.8% of students answered the question correctly. Questions 7,8,10, and 12 from the assessment were used on the questionnaire and covered Algebra and Functions. 46.9% of students correctly answered question 7, 45.1% of students correctly answered question 8, 71.4% of students correctly answered question 10, and 62.5% of students correctly answered question 12. Questions 13,14, and 17 from the assessment were used on the questionnaire and covered Geometry and Measurement. 55.8% of students correctly answered question 13, 80.9% of students correctly answered question 14, and 58.5% of students correctly answered question 17. Question 19 from the assessment was used on the questionnaire and covered Data Analysis. 63.3% of students correctly answered question 19.

5. Empirical Strategy

1. Impact of Calculator Use in Open Ended Test Framework and Impact of Multiple Choice

To analyze the impact of each type of calculator (basic, graphing, and own calculator) in the open ended framework as well as the impact of using multiple choice instead of open ended framework, we use a difference in difference framework as follows:

$y_i = \alpha + \beta\tau_i + \sum_{j=1}^{4}(\delta_j T_i j) + \sum_{j=1}^{4}(\gamma_j T_i j) * \tau_i + \epsilon_i$

where i refers to student i, j refers to treatment group j (1 through 4 listed below), y_i is the score of student i (total out of 10 points and 0,1 for each question evaluated separately), $T_i 1 = 1$ if student i received multiple choice framework, $T_i 2 = 1$ if student i received a basic calculator, $T_i 3 = 1$ if student i received a graphing calculator, $T_i 4 = 1$ if student i was allowed to use his/her own calculator, where the control group is students that received the open ended test framework without the use of a calculator.

2. Impact of Calculator Use in Multiple Choice Framework

To analyze if there are any interactive impacts of using a calculator with multiple choice, we estimate the following equation:

$y_i = \alpha + \beta\tau_i + \sum_{j=2}^{4}(\delta_j T_i j) + \sum_{j=2}^{4}(\gamma_j T_i j) * \tau_i + \epsilon_i$

where i refers to student i, j refers to treatment group j (2 through 4 listed below), y_i is the score of student i (total out of 10 points and 0,1 for each question evaluated separately), $T_i 2 = 1$ if student i received a basic calculator, $T_i 3 = 1$ if student i received a graphing calculator, $T_i 4 = 1$ if student i was allowed to use his/her own calculator, where the control group is students that received the multiple choice test framework without the use of a calculator.

6. Results

Results for the impact of calculator use and multiple choice format as well as the interaction between the two on students average total scores can be seen in Table 5 below. We find that allowing a student to use any calculator in an open ended test framework improves a student's score by .5-.6 of a point (or 5-6 percentage points) significant at the 1% level.[4] We also find that introducing multiple choice improves a students score by .9 of a point (or 9 percentage points), again significant at the 1% level. There is a statistically significant increase in a student's test score by an additional .5 points when introducing a graphing calculator in addition to the multiple choice framework (implying students that received multiple choice and a graphing calculator score 1.4 points higher- i.e. increasing their score by 14 percentage points-in comparison to those that received open ended framework without the use of a calculator). There is no statistically significant effect from introducing other forms of calculator in the multiple choice framework.[5] The control group improved their score by .2 points (or 2 percentage points) implying that simply seeing the exam an additional time jogged students memories and yielded an improvement on their overall score.

Results for the impact of calculator use and multiple choice format as well as the interaction between the two on the percentage of students that answered each question correctly can be seen in Table 6 and 7 below. The most notable and interesting results are discussed here. Of the Numerics and Operations questions used in the questionnaire, question 1 asked students to calculate percentage change. Results from Table 6 demonstrate that giving students a basic calculator improves the percentage of students that answer correctly by 6.6 percentage

[4]There is no statistically significant difference between each type of calculator used.

[5]Remember that only half of students brought their own graphing calculator implying that results where students use multiple choice and are given a graphing calculator are likely the most comparable to the SAT Math as almost all students bring graphing calculators to the SAT Math.

points with no statistically significant effect from any other type of calculator. Giving students multiple choice improves the percentage of students that answer correctly by 8.5 percentage points. Originally, 90% of students correctly answered this question, therefore giving students multiple choice increases this to 98.5%. There is not statistically significant interactive impact from giving students a calculator in either test framework.

Of the Algebra and Functions questions used in the questionnaire, questions 3 and 5 are both algebraic word problems that require students to set up two equations and solve for two unknowns. We find that in question 3, giving students a calculator improves the percentage of students that answer correctly by 12-23 percentage points (with no statistically significant difference for each type of calculator). Originally, only 47% of students correctly answered this question, therefore giving students a calculator increases this to 59-70%. Giving students multiple choice improves the percentage of students that answer correctly by 21 percentage points, therefore increasing the percentage that correctly answer the problem from 47% to 68%. The impact from a calculator and multiple choice are similar. Analyzing the interaction between test framework and calculator use one can see that adding a graphing calculator to the multiple choice improves the percentage of students that correctly answer the problem by an additional 18 percentage points, thereby increasing the percentage of students that correctly answer the question from 68% to 86%. There is no significant effect from any other type of calculator.

For question 5, giving students a calculator improves the percentage of students that answer correctly by 7-10 percentage points (again with no statistically significant difference for each type of calculator). Originally, 71% of students correctly answered this question, therefore giving students a calculator increases this to 78-81%. Giving students multiple choice improves the percentage of students that answer correctly by 8 percentage points, therefore increasing the percentage that correctly answer the problem from 71% to 79%. Again the impact from giving student a calculator versus multiple choice are close. Analyzing the interaction between test framework and calculator use one can see that adding a graphing calculator to the multiple choice improves the percentage of students that correctly answer the problem by an additional 8 percentage points, thereby increasing the percentage of students that correctly answer the question from 79% to 87%. There is no significant effect from any other type of calculator. For question 5, by simply giving students the exam again there is an increase in the percentage of students that correctly answer the question by 4.5 percentage points.

The difference between results for questions 3 and 5 suggests a increasing return to calculators and multiple choice framework with increasing difficulty level. Question 3 was more difficult for students to answer and we can see a much larger impact of calculators and multiple choice for this question.

Of the Algebra and Functions questions used in the questionnaire question 6 asks students to calculate the minimum value of a function. Results from Table 7 show that giving students a calculator on this question improves the percentage of students that correctly answer the problem by 5-16 percentage points (with no statistically significant difference for each type of calculator). Originally, 63% of students correctly answered this question, therefore giving students a calculator increases this to 68-79%. Giving students multiple choice improves the percentage of students that answer correctly by 10 percentage points, implying that the percentage of students that answer correctly would increase from 63% to 73%. There was no significant impact from adding a calculator to the multiple choice framework. Again there is little difference in the impact from having a calculator versus having multiple choice.

Of the Geometry and Measurement questions asked on the questionnaire, there was no statistically significant impact from giving students multiple choice or a calculator for question 7 where students are required to calculate the area of a pentagon. Considering that only 56% of students correctly answered this question this is an interesting result. For question 8, where students are asked for the slope of the line drawn, giving students the use of their own calculator improves the percentage of students that answer correctly by 8 percentage points, with no significant effect from any other calculators. This result could be due to the fact that students are familiar with the graphing properties of their own calculator. Giving students multiple choice test framework improves the percentage that answer the question correctly by 9 percentage points. Given that originally 81% answered the question correctly, this increases the percentage that answer correctly to 90%. There is no statistically significant interactive impact from calculator and multiple choice framework.

Lastly, in question 10, which covers the data analysis topic, we find that giving students multiple choice framework increased the percentage of students that answered correctly from 63% to 81% (an increase of 18 percentage points). There was no statistically significant impact from having a calculator and no interactive impact from having a calculator and multiple choice. By simply giving the students the exam again, there is an increase in the percentage of students that correctly answer the question by 9 percentage points.

This final question brings into question a serious qualm that people have with multiple choice. In this question, students were asked for the number of people represented by a share of the circle graph. The problem students had in the open ended framework was reading the question thoroughly. When originally answering the question on the assessment, many students simply put down the percentage of people represented by the share of the circle graph instead of the number of people. The share was .5 or 50% and students were required to find the number of people represented by that share when given the population. However, of the 4 possible choices in the multiple choice framework, 50 was not given as a choice. For students that received multiple choice, once they realized their answer was not there, they re-read the question. In economics, if you interpret a percentage as a value (for instance maybe an elasticity as a quantity or price) the answer is completely incorrect. In economics interpretation is key and giving students multiple choice does not allow for the testing of this interpretation.

7. Conclusion

To conclude, we have verified our hypothesis that test format and calculator use are key contributors as to why the Algebra I Assessment is a better predictor of students performance in principles of economics than the SAT Math. This implies that if one is to implement a similar assessment he/she should be wary of using a multiple choice and/or allowing for the use of any calculator.

We find that calculator and test format have an impact on students overall scores, with an interactive impact from adding a graphing calculator to the multiple choice framework. As mentioned before this is likely the closest comparison group to the SAT Math.

We find that the impact of calculator use and test format varies depending on the type of question asked. For instance, in questions relating to Algebra and Functions where a word problem is used and students are expected to write out the functional form and solve for the two unknowns, calculator use and multiple choice alone have similar impacts, while adding a graphing calculator to the multiple choice framework has a large impact on the percentage of students that correctly answer the problem. Whereas in the case of the percentage change problem and data analysis problem, multiple choice had the largest impact, likely because students that have a sense of numeracy can analyze the possible answers and determine which answers are most likely correct when given multiple potential answers.

A basic math assessment that is open-ended and does not allow use of a calculator is a better predictor of student performance in a Principles of Economics course than SAT Math scores. Administering such an assessment will allow faculty to determine which students do not have the math ability to remain in the course and which students could take the course with some remedial help. Such assessment also gives students the correct signal about the rigor of the course. We emphasize the importance of making the assessment score a part of the course grade in order for students to take the math preparation seriously.

REFERENCES

Allwine, Melani, and Irene Foster. 2014. "The Role of Algebra I Assessment in Improving Student Performance in Economics."

Ballard, Charles, and Marianne Johnson. 2004. "Basic Math Skills and Performacne in an Introductory Economics Class." *Journal of Economic Education*, 35(1): 3–23.

Becker, William, and Carol Johnston. 1999. "The Relationship between Multiple Choice and Essay Response Questions in Assessing Economics Understanding." *Economic Record*, 75: 348–357.

Bridgeman, Brent, and Maria Potenza. 1998. "Effects of an On-Screen versus Bring-Your-Own Calculator Policy on Performance on the Computerized SAT I: Reasoning Test in Mathematics." *Presented at the Annual Meeting of the National Council on Measurement in Education (NCME).*

Bridgeman, Brent, Anne Harvey, and James Braswell. 1995. "Effects of Calculator Use on Scores on a Test of Mathematical Reasonin." *Journal of Educational Measurement*, 32(4): 323–340.

Lagerlof, Johan, and Andrew Seltzer. 2009. "The effects of remedial mathematics on the learning of economics: Evidence from a natural experiment." *The Journal of Economic Education*, 40(2): 115–137.

Mallik, Girijasankar, and John Lodewijks. 2010. "Student performance in a large first year economics subject: Which variables are significant?" *Economic Papers: A journal of applied economics and policy*, 29(1): 80–86.

Owen, Ann. 2012. "32 Student characteristics, behavior, and performance in economics classes." *International Handbook on Teaching and Learning Economics.*

Pozo, Susan, and Charles Stull. 2006. "Requiring a math skills unit: Results of a randomized experiment." *The American Economic Review*, 437–441.

Table 1—Algebra I Assessment Questions and Relationship to Economics

Question Number	Corresponding SAT Math Section	Representation as % in Algebra I Assessment	Representation as % in SAT Math	Relationship to Economics
1	Numerics and Operations	20%	20-25%	Taxes
2	Numerics and Operations	20%	20-25%	Growth rates; elasticity
3	Numerics and Operations	20%	20-25%	Setting up demand or supply from a word problem
4	Numerics and Operations	20%	20-25%	Comparative statics; solving for the equilibrium
5	Algebra and Functions	40%	35-40%	Solving for the equilibrium; setting up demand or supply from a word problem
6	Algebra and Functions	40%	35-40%	Solving for the equilibrium; 2 equation and 2 unknowns
7	Algebra and Functions	40%	35-40%	Solving for the equilibrium; 2 equation and 2 unknowns
8	Algebra and Functions	40%	35-40%	Solving for demand/inverse demand
9	Algebra and Functions	40%	35-40%	Cost function
10	Algebra and Functions	40%	35-40%	Solving for the equilibrium; 2 equations and 2 unknowns
11	Algebra and Functions	40%	35-40%	Supply equations
12	Algebra and Functions	40%	35-40%	Solving for the minimum
13	Geometry and Measurement	30%	25-30%	Consumer/produce surplus
14	Geometry and Measurement	30%	25-30%	Demand and supply
15	Geometry and Measurement	30%	25-30%	Demand and supply
16	Geometry and Measurement	30%	25-30%	Consumer surplus
17	Geometry and Measurement	30%	25-30%	Change in consumer surplus, producer surplus, and national welfare
18	Geometry and Measurement	30%	25-30%	Demand and supply
19	Data Analysis	10%	10-15%	Using circle graphs to interpret data
20	Data Analysis	10%	10-15%	Cost tables

Table 2—Differences in share of students that did not consent to partake in the experiment across treatment and control groups.

Mean in Open Ended No Calculator	*Difference between* Multiple Choice No Calculator *and* Open Ended No Calculator	*Difference between [....] and* Open Ended No Calculator			*Difference between [....] and* Multiple Choice No Calculator		
		Open Ended with basic calculator	Open Ended with graphing calculator	Open Ended with own calculator	Multiple Choice with basic calculator	Multiple Choice with graphing calculator	Multiple Choice with own calculator
1	2	3	4	5	6	7	8
0.056	0.026	0.033	0.03	0.017	-0.023	-0.034	-0.024
	(0.0307)	(0.033)	(0.033)	(0.031)	(0.034)	(0.033)	(0.034)

Table 3—Experiment Validity

	Mean in Open Ended without calculator	*Difference between* Multiple Choice without calculator *and*	*Difference between [....] and* Open Ended without calculator			*Difference between [....] and* Multiple Choice without calculator		
		Open Ended without calculator	Open Ended with basic calculator	Open Ended with graphing calculator	Open Ended with own calculator	Multiple Choice with basic calculator	Multiple Choice with graphing calculator	Multiple Choice with own calculator
	1	2	3	4	5	6	7	8
SAT Math	653	8.78	8.91	1.92	3.18	-7.37	-14.09*	-13
		(9.5)	(10.15)	(10.27)	(10.55)	(10.03)	(10.2)	(10.38)
ACT Math	28.52	-0.126	0.182	0.124	0.288	-0.295	-0.211	-0.35
		(0.83)	(0.777)	(0.85)	(0.75)	(0.863)	(0.809)	(0.78)
% female students	0.53	0.048	0.009	0.067	0.029	0.054	0.053	0.029
		(0.061)	(0.065)	(0.065)	(0.064)	(0.067)	(0.066)	(0.066)
% in private high school	0.468	0.05	0.052	0.035	.1*	-0.042	-0.078	-0.078
		(0.06)	(0.065)	(0.065)	(0.063)	(.064)	(0.064)	(0.064)

Table 4—Assessment Scores for Experiment Participants

Questionnaire Question #	Assessment Question #	*Mean on the Assessment*
1	2	0.896
2	4	0.968
3	7	0.469
4	8	0.451
5	10	0.714
6	12	0.625
7	13	0.558
8	14	0.809
9	17	0.585
10	19	0.633
	Total Score (out of 10)	7.7
	Total Score (out of 20)	13.4
	Fail (=1 if fail) (out of 10)	0.615
	Fail (=1 if fail) (out of 20)	0.66

Table 5—Results for Total Score

		Impact of Multiple Choice and Impact of Calculator use in Open Ended Framework	*Impact of Calculator use in Multiple Choice Framework*
		A	B
Multiple Choice	γ_1	.883***	
		(0.221)	
Basic Calculator	γ_2	0.602***	236
		(.148)	(0.213)
Graphing Calculator	γ_3	.471***	0.46***
		(0.159)	(0.137)
Own Calculator	γ_4	.519***	0.218
		(0.154)	(0.152)
	β	.224*	1.1***
		(.112)	(0.128)
	N	544	412
	R^2	0.08	0.014

Table 6—Results for Questions 1 through 5

		Question 1		Question 2		Question 3		Question 4		Question 5	
		Impact of Multiple Choice and Impact of Calculator use in Open Ended Framework	*Impact of Calculator use in Multiple Choice Framework*	*Impact of Multiple Choice and Impact of Calculator use in Open Ended Framework*	*Impact of Calculator use in Multiple Choice Framework*	*Impact of Multiple Choice and Impact of Calculator use in Open Ended Framework*	*Impact of Calculator use in Multiple Choice Framework*	*Impact of Multiple Choice and Impact of Calculator use in Open Ended Framework*	*Impact of Calculator use in Multiple Choice Framework*	*Impact of Multiple Choice and Impact of Calculator use in Open Ended Framework*	*Impact of Calculator use in Multiple Choice Framework*
		1A	1B	2A	2B	3A	3B	4A	4B	5A	5B
Multiple Choice	δ_1	.085***		0.016		0.208***		0.046		.078***	
		(0.021)		(0.025)		(0.067)		(0.05)		(0.026)	
Basic Calculator	δ_2	.066**	.053	0.011	-0.016	0.231***	0.057	-0.034	0.035	0.075*	0.035
		(0.026)	(.037)	(0.01)	(0.025)	(0.051)	(0.11)	(0.051)	(0.054)	(0.039)	(0.029)
Graphing Calculator	δ_3	0.022	.038	0.021**	-0.027	.202***	0.18**	-0.002	0.062	0.102***	0.081**
		(.029)	(.033)	(0.009)	(0.034)	(0.04)	(0.082)	(0.035)	(0.061)	(0.031)	(0.038)
Own Calculator	δ_4	.011	.031	0.02**	0.015	.129**	0.041	0.005	0.002	.094*	0.011
		(.021)	(.031)	(0.009)	(0.029)	(0.048)	(0.075)	(0.033)	(0.041)	(0.06)	(0.05)
	α	-0.011	0.074	0	0.016	0.03	0.238	0.045	.091**	.045***	.123***
		(.016)	(0.023)		(0.025)	(.034)	(0.048)	(.033)	(0.039)	(.012)	(0.02)
	N	544	412	544	412	544	412	544	412	544	412
	R^2	0.03	0.013	0.005	0.014	0.05	0.019	0.01	0.003	0.014	0.006

Table 7—Results for Questions 6 through 10

		Question 6		Question 7		Question 8		Question 9		Question 10	
		Impact of Multiple Choice and Impact of Calculator use in Open Ended Framework	*Impact of Calculator use in Multiple Choice Framework*	*Impact of Multiple Choice and Impact of Calculator use in Open Ended Framework*	*Impact of Calculator use in Multiple Choice Framework*	*Impact of Multiple Choice and Impact of Calculator use in Open Ended Framework*	*Impact of Calculator use in Multiple Choice Framework*	*Impact of Multiple Choice and Impact of Calculator use in Open Ended Framework*	*Impact of Calculator use in Multiple Choice Framework*	*Impact of Multiple Choice and Impact of Calculator use in Open Ended Framework*	*Impact of Calculator use in Multiple Choice Framework*
		6A	6B	7A	7B	8A	8B	9A	9B	10A	10B
Multiple Choice	δ_1	.106***		0.007		0.089**		0.066		0.181***	
		(0.016)		(0.059)		(0.034)		(0.095)		(0.018)	
Basic Calculator	δ_2	0.046***	0.038	0.09	0.014	0.011	-0.03	0.106**	0.097	0.008	0.056
		(0.015)	(0.028)	(0.055)	(0.039)	(0.025)	(0.02)	(0.042)	(0.063)	(0.045)	(0.056)
Graphing Calculator	δ_3	0.118**	0.044	0.005	-0.088	0.032	0.05	0.037	0.13*	-0.068	-0.005
		(0.045)	(0.028)	(0.045)	(0.084)	(0.024)	(0.039)	(0.048)	(0.067)	(0.048)	(0.02)
Own Calculator	δ_4	.161*	0.108	-0.014	-0.001	0.081**	0.025	-0.0004	-0.001	0.029	-0.002
		(0.09)	(0.105)	(0.046)	(0.046)	(0.028)	(0.038)	(0.044)	(0.068)	(0.025)	(0.067)
	α	-0.007	0.098	-0.052	-0.045	-0.011	.078***	.09**	0.156	.09***	.27***
		(.012)	(0.022)	(.039)	(0.031)	(.02)	(0.016)	(.037)	(0.063)	(.018)	(0.016)
	N	544	412	544	412	544	411	544	412	544	412
	R^2	0.036	0.011	0.007	0.004	0.021	0.007	0.02	0.015	0.06	0.003

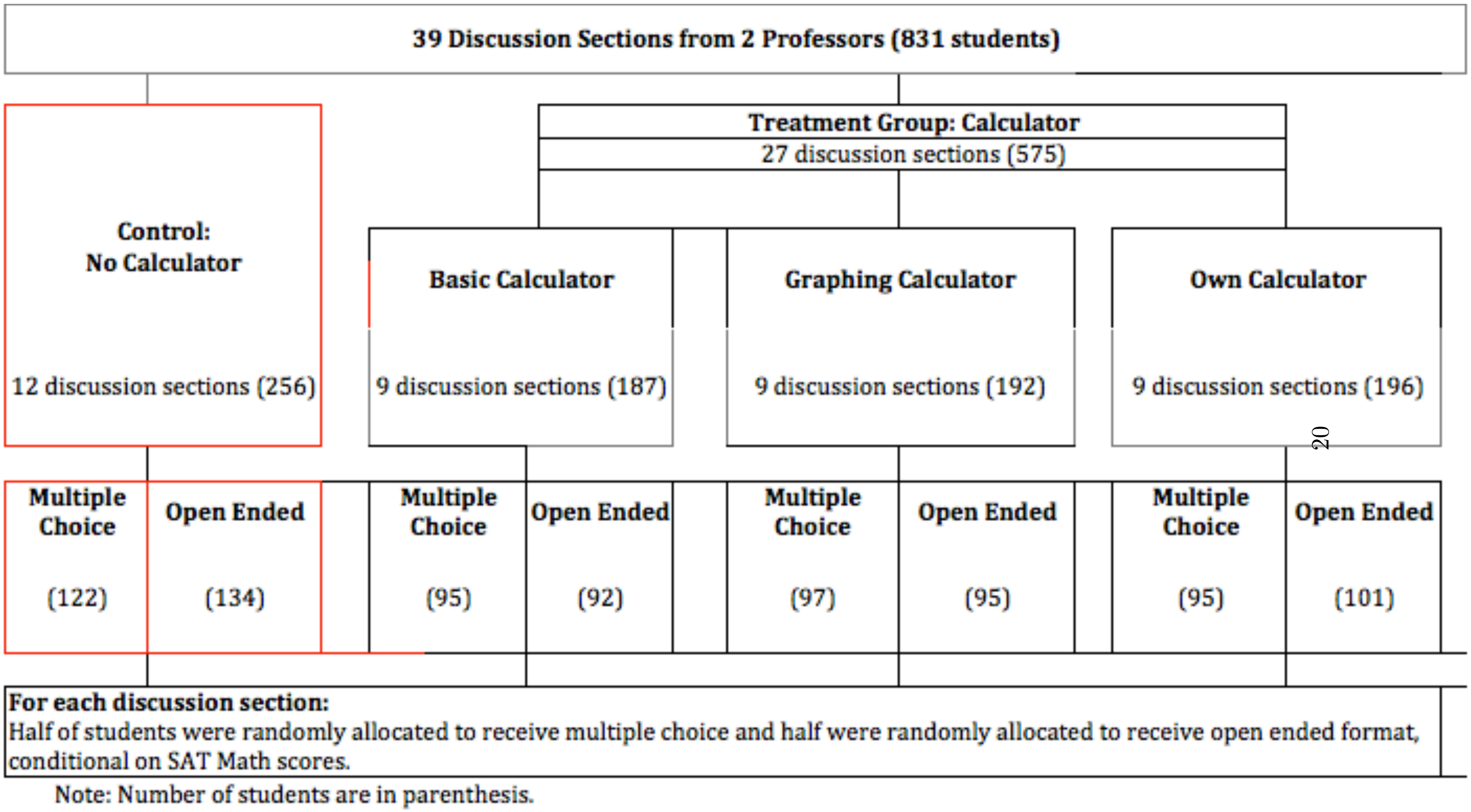

Figure 1. Experiment Design