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

Melanie Allwine and Irene R. Foster[1]
Department of Economics, The George Washington University

December 2013

Preliminary version – please do not cite without authors' permission[2]

Abstract

This research examines if a simple Algebra I assessment (no calculators allowed and open-ended format) administered by the Principles of Economics faculty at this institution is capturing some weakness in students' basic math ability that the SAT Math test (calculator use allowed and mostly multiple-choice format) is not. Our hypothesis is that the use of a calculator and a multiple-choice format may be allowing students to answer questions correctly they would otherwise not be able to work through and solve. In other words, students may know how to compute answers with calculators or reverse engineer multiple-choice questions while failing to understand underlying math concepts.

This paper presents results from primary experimental data collected Fall 2013 on 1400 students registered for Principles of Economics for which the prerequisite is Algebra I. The dataset contains Algebra I assessment results by type of question (with and without the use of a calculator; and either multiple choice or open-ended format) and other student characteristics including SAT Math scores. Assessment questions were identical to those on recent SAT Math tests.

We identify whether calculator use or test format has a greater effect overall on speed and performance on tests. We identify questions students are able to answer correctly with or without a calculator, and multiple choice or open-ended format. We then identify questions students are unable to answer without a calculator or if a test has an open-ended format, and whether type of calculator makes a difference. We compare performance on such questions to see which of these math skills are critical to the study of economics. Results are presented by student characteristics. Finally, we discuss how to develop assessments to test math skills needed for Principles of Economics.

[1] Corresponding author: Irene R. Foster – Email: fosterir@gwu.edu
[2] Institutional Review Board permission was granted to conduct this research at The George Washington University (IRB #081330).

1. Introduction

Each year, students' college admissions are based in part on their SAT Math scores. The expectation of Economics faculty is that if students have been admitted into college, they should have the math skills necessary to navigate a Principles of Economics course. However, anecdotal evidence from faculty suggests that this is not the case. Evidence from research in this area suggests that many students fail to comprehend foundational material in economics due to gaps in their understanding of basic math concepts (Mallik and Lodewijks 2010; Owen 2012) and that in addition to grades in advanced math classes and strong standardized test scores, a student's score on a math quiz given at the beginning of an economics class has a statistically significant positive impact on performance in the course (Arnold and Straten 2012, Ballard and Johnson 2004, Benedict and Hoag 2002, Benedict and Hoag 2012, Schuhmann et al 2005).

To address this issue, Principles faculty at this institution administer a common Algebra I assessment at the start of each Fall semester. The questions are simple and cover only those skills required for a Principles class – percentage change, ratios, proportions, fractions and decimals, order of operations, place value, the area of a triangle, simple exponents and the graph of a straight line. Students are informed two weeks before class that an Algebra I assessment worth 10% of the course grade will be administered in Week One. They are pointed towards learning resources and told that calculators will not be allowed on the assessment. Any student who does not achieve an 80% (indicating mastery) on the first try has three other chances to pass the assessment during the first part of the semester. Math reviews are offered every other week in between assessment opportunities. Given that it is 10% of their course grade, students take passing the assessment quite seriously. Given the importance of remedial mathematics for learning economics (Lagerlöf and Seltzer 2009), students who do not pass the assessment even after four attempts are strongly encouraged to take the class only after they have mastered the math prerequisite. In this manner, faculty can maintain the rigor of the class while ensuring that students are not falling behind.

Allwine and Foster (2013) looked at primary data collected on 1361 Principles of Economics students to understand whether Algebra I assessment scores (from the first attempt) predict course performance. Similar to prior research in this area (Ballard and Johnson 2004, Cohn et al 2001, Pozo and Stull 2006), results indicate that assessment scores do predict course performance, and that students are more motivated to learn the math (and therefore perform better in the course) when the assessment score is a significant part of their course grade. Interestingly, however, despite the fact that the SAT Math and Algebra I Assessment covered some of the same material, Allwine and Foster (2013) found that students' SAT Math scores had little explanatory power with respect to students' performance in Principles of Economics. The paper suggested some reasons for the difference. It could be that the SAT Math test has broader coverage than the Algebra I Assessment that only covers math concepts specifically related to the Principles of Economics course;

questions on the two tests may be worded very differently; the SAT Math test allows calculator use while the assessment does not; and test format is mostly multiple choice on the SAT while it is open-ended on the assessment.

Given the importance of SAT Math scores in admissions decisions, this paper tests the hypothesis that calculator use and a multiple choice test format mask weaknesses in students' math ability that affect their performance in a Principles class. Primary experimental data was collected in the Fall of 2013 on 1400 students registered for Principles of Economics for which the prerequisite is Algebra I. All students first took a simple 20 question open-ended Algebra I Assessment without the use of a calculator. Assessment questions were identical to those on recent SAT Math tests. Students were then divided into experimental and control groups for two treatments – test format (open-ended or multiple-choice) and calculator (no calculator, basic calculator, graphing calculator[3] or own calculator) – and asked to complete another math assessment which contained half the questions from the previous assessment.

Section 2 describes previous research in this area. In Section 3 we present our overall assessment results. In section 4 we analyze the explanatory power of students' SAT Math scores in predicting their performance on the Algebra I assessment. In section 5 we use students' scores on the SAT Math test as well as their performance on the Algebra I assessment to predict performance on the final exam in the Principles class. In Section 6 we present our conclusions.

2. Literature Review

Why don't SAT Math scores have the same predictive power as scores on the Algebra I Assessment? This paper tests the hypothesis that it is due to calculator use and multiple-choice format on the SAT Math test.

Calculator Use

Prior research suggest that the use of a calculator on the SAT may be allowing students to answer questions they otherwise would not be able to think through and solve. Across all precollege grade and ability levels, students using calculators possess better attitude toward mathematics and a better self-concept in mathematics than students not using calculators, although sustained calculator use seems to hinder development of basic problem-solving skills in average students (Hembree and Dessart 1986). Bridgeman, Harvey and Braswell (1995) note that students note that students benefit from using calculators but fail to understand underlying concepts. Their speed was not significantly affected by whether or not they used calculators.

[3] TI-83 graphing calculator.

Roberts (1980) indicated that calculator use allows for computational benefits but not conceptual benefits. Ellington (2003) showed that the computational benefits of calculators may not necessarily allow students to retain and transfer math skills to other subjects. Would students perform better on math tests if allowed to use their own calculator? Prior research shows there are no performance advantages with students using their own calculators (Hanson et al 2001, Bridgeman and Potenza 1998)

Test Format

Work done by Erickan el al (1998) shows that constructed-response or open-ended questions give a better indication of a low ability and high ability than multiple-choice questions. Becker and Johnston (1999) argue that multiple-choice and open-ended questions evaluate different dimensions of knowledge and so both could be used for testing in economics.

Although using a programming model instead of a purely mathematical test, Simkin and Kuechler (2005) show that it is more difficult to compose multiple-choice (MC) items to reach the highest knowledge levels achieved by open-ended or constructed-response (CR) questions. This could therefore show that MC questions are easier and do not fully test a student's knowledge compared to CR questions. Later work (Kuechler and Simkin 2010) shows weak correlation between MC and CR portions of entry level programming tests. The multiple-choice format of the SAT Math test allows students to answer questions correctly without fully understanding what they are doing (Becker and Johnston 1999; Rebeck and Asaarta 2012).

3. The Algebra I Assessment

The layout of the Algebra I Assessment was as follows: 20 questions were chosen from recent SAT Math tests 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 in Table 1. Questions were chosen that were simple and could be computed easily without the use of a calculator. The questions chosen covered the same topics as the SAT Math test in the same proportions: 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 attached.

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 questions 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.

4. Experiment

As soon as all students were given the Algebra I Assessment for the course, 902 students[4] 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.[5] This sample of students included 13 teaching assistants and 39 discussion sections, where each teaching assistant has 3 discussion sections with approximately 25 students in each. 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 2 demonstrates that there is no statistically significant difference between the percentage of students that did not consent in treatment and control groups.

Figure 1 is a representation of the experimental design used. Teaching assistants were randomly assigned to a calculator treatment or control conditional on students' average SAT Math score.[6] 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 calculators used for the assessment had only addition, subtraction, multiplication, and division functions. Exponential and square root functions were not available. Students in the treatment where they were using their own calculators were asked to report if the calculator was a basic calculator or graphing calculator. According to student self-reports, 40% of students used a basic calculator, 56% used a graphing calculator, and 4% used no calculator. Although all students were advised to bring a calculator to the assessment, it is likely that many forgot. 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.

---

[4] Of the 1144 students enrolled in the large principles of microeconomics courses 242 were dropped from the sample.

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

[6] Randomization was first done at the teaching assistant level to ensure that treatment was followed as it would have caused students unneeded stress and would have been difficult to enforce if randomization was done by the discussion section or student.

To further test any interactions between calculator use and test format, we randomize 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 3 below, 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).

5. Assessment Results for 831 students that participated in the experiment

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 66% of the 20 questions asked correctly. This led to a failure rate of 66% on the first attempt at the assessment.

Table 4 shows the average total score on the assessment for these students if only the 10 questions used in the Questionnaire had been used 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 also 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.

6. Model

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:

$$yiQ-yi\,A=\alpha+j=14\delta jTj+\varepsilon i \quad (1)$$

where *i* refers to student *i*, *j* refers to treatment group *j* (1 through 4 listed below), y_{iQ} is the score of student *i* on the questionnaire (total and for each question separately), y_{iA} is the score of student *i* on the assessment (total out of 10 points and for each question separately), T_1=1 if the student received multiple choice test framework, T_2=1 if the student received a basic calculator, T_3=1 if the student received a graphing calculator, and T_4=1 if the student 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.

Impact of Calculator Use in Multiple Choice Framework

This analysis will allow us to see if there is any interactive impact of using a calculator with multiple choice. For instance, it may be the case that students can figure out an answer more easily if they have a calculator along with multiple choice whereas having a calculator in and of itself may not be helpful.

To analyze this interaction effect again we use a difference in difference framework as follows:

$$yiQ-yi\,A=\alpha+j=24\delta jTj+\varepsilon i \quad (2)$$

Where *i* refers to student *i*, *j* refers to treatment group *j* (2 through 4 listed below), y_{iQ} is the score of student *i* on the questionnaire (total and for each question separately), y_{iA} is the score of student *i* on the assessment (total out of 10 points and for each question separately), T_2=1 if the student received a basic calculator, T_3=1 if the student received a graphing calculator, and T_4=1 if the student was allowed to use his/her own calculator.

7. 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 a calculator in an open-ended test framework improves a student's score by 0.5-0.6 of a point (or 5-6 percentage points) *significant at the 1% level*. We also find that introducing multiple-choice improves a students score by 0.9 of a point (or 9 percentage points). There is a statistically significant increase in a student's test score by an additional 0.5 points when introducing a graphing calculator in additional 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 the open-ended framework without the use of a calculator). There is no statistically significant effect from introducing other forms of calculator.[7] The control group improved their score by 0.2 points (or 2 percentage points) implying that simply seeing the questions for the second 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 Part I and Part II 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 Part I demonstrate that giving students a basic calculator improves the percentage of students that answer correctly by 6.6 percentage 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 along with the multiple-choice.

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. Originally, only 47% of students correctly answered this question, therefore giving students a calculator increases this to 59-70%. Giving students multiple choice improve 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.

---

[7] Remember that only half of students brought their own graphing calculator implying that results where students use MC 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.

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. 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 of giving students 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.

That question 3 is more difficult for students to answer without the use of multiple choice and a graphing calculator is an interesting phenomenon. We do not believe this is due to the difficulty of the computation, because the effect would have been significant and positive for all calculator treatment groups. It is likely due to the wording of the problems. It may be easier for students to relate to question 5 and therefore be able to solve it intuitively using their mathematical reasoning.

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 6 Part II show that giving students a calculator on this question improves the percentage of students that correctly answer the problem by 5-16 percentage points. 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, much of 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 below the age of 40 instead of the number. However, of the 4 possible choices, 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 question is completely incorrect. In economics interpretation is key and giving students multiple choice does not allow for the testing of this interpretation.

8. 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 to assess students' readiness for economics, he/she should not use a multiple choice or open-ended format.

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 test takers.

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.

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 remina 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, Melanie and Irene R. Foster, “Algebra I Assessment and Student Performance in Principles of Economics.” The George Washington University, unpublished working paper (2013).

Arnold, Ivo J. M., and Jerry T. Straten. "Motivation and math skills as determinants of first-year performance in economics." *The Journal of Economic Education* 43.1 (2012): 33-47.

Ballard, Charles L., and Marianne F. Johnson. "Basic math skills and performance in an introductory economics class." *The Journal of Economic Education* 35.1 (2004): 3-23.

Becker, W. E. and Johnston, C. “The Relationship between Multiple Choice and Essay Response Questions in Assessing Economics Understanding.” *Economic Record* (1999), 75: 348–357.

Benedict, Mary Ellen, and John Hoag. "Who's afraid of their economics classes? Why are students apprehensive about introductory economics courses? An empirical investigation." *The American Economist* (2002): 31-44.

Benedict, Mary Ellen, and John Hoag. "Factors Influencing Performance in Economics: Graphs and Quantitative Usage." *International Handbook on Teaching and Learning Economics* (2012): 334-340.

Bridgeman, Brent, Anne Harvey, and James Braswell. "Effects of calculator use on scores on a test of mathematical reasoning." *Journal of Educational Measurement* 32.4 (1995): 323-340.

Bridgeman, B. and M. Potenza. ”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) 1998.
Retrieved October 28, 2013 from http://www.editlib.org/p/83725.

Cohn, Elchanan, et al. "Do graphs promote learning in principles of economics?" *The Journal of Economic Education* 32.4 (2001): 299-310.

Ellington, Aimee J. “A meta-analysis of the effects of calculators on students’ achievement and attitude levels in precollege mathematics classes.” *Journal for Research in Mathematics Education,* Vol 34, No. 5 (November 2003), 433-463.

Ercikan, K., Sehwarz, R. D., Julian, M. W., Burket, G. R., Weber, M. M. and Link, V. "Calibration and Scoring of Tests With Multiple-Choice and Constructed-Response Item Types." *Journal of Educational Measurement* (1998), 35: 137–154.

Hanson, Kaaren, Bonny Brown, Roger Levine and Teresa Garcia. "Should standard calculators be provided in testing situations? An investigation of performance and preference differences." *Applied Measurement in Education*, Vol. 14, Issue 1 (2001).

Hembree, Ray and Donald J. Dessart. "Effects of hand-held calculators in precollege mathematics education." *Journal for Research in Mathematics Education*, Vol. 17, No. 2 (March 1986): 83-99.

Kuechler, W. L. and Simkin, M. G. "Why Is Performance on Multiple-Choice Tests and Constructed-Response Tests Not More Closely Related?" *Theory and an Empirical Test. Decision Sciences Journal of Innovative Education*, (2010), 8: 55–73.

Lagerlöf, Johan N. M, and Andrew J. Seltzer. "The effects of remedial mathematics on the learning of economics: Evidence from a natural experiment." *The Journal of Economic Education* 40.2 (2009): 115-137.

Mallik, Girijasankar, and John Lodewijks. "Student performance in a large first year economics subject: Which variables are significant?." *Economic Papers: A journal of applied economics and policy* 29.1 (2010): 80-86.

Owen, Ann L. "32 Student characteristics, behavior, and performance in economics classes." *International Handbook on Teaching and Learning Economics* (2012): 341.

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

Rebeck, Ken and Carlos Asarta. "Methods of Assessment in the College Economics Course." *International Handbook on Teaching and Learning Economics* (2012): 177-187.

Roberts, Dennis M., "The impact of electronic calculators on educational performance." Review of Educational Research, (Spring 1980), 50: 71-98.

Schuhmann, P., K. McGoldrick and R. Burrus. "Student quantitative literacy: Importance, measurement, and correlation with economic literacy." *The American Economist* (2005). 49(4); 49-65.

Simkin G. Mark, and Kuechler L. William, "Multiple choice tests and student understanding: What is the connection?" (2005). *Decision Sciences Journal of Innovative Education*. Volume 3-1.

Table 1

Question	Corresponding SAT Math Section	Representation as % in Algebra I Assessment	Representation as % in SAT Math	Relationship to Economics
1	Numerics and Operations	20.00%	20-25%	taxes
2	Numerics and Operations	20.00%	20-25%	growth rates; elasticity
3	Numerics and Operations	20.00%	20-25%	setting up demand or supply from word problem
4	Numerics and Operations	20.00%	20-25%	comparative statics; Solving for equilibrium
5	Algebra and Functions	40.00%	35-40%	Solving for equilibrium price or quantity; setting up demand or supply from word problem
6	Algebra and Functions	40.00%	35-40%	Solving for equilibrium ; 2 equations and 2 unknowns
7	Algebra and Functions	40.00%	35-40%	Solving for equilibrium; 2 equations and 2 unknowns
8	Algebra and Functions	40.00%	35-40%	Solving for demand/ inverse demand
9	Algebra and Functions	40.00%	35-40%	Cost function
10	Algebra and Functions	40.00%	35-40%	Solving for equilibrium; 2 equations and 2 unknowns
11	Algebra and Functions	40.00%	35-40%	Understanding supply equations
12	Algebra and Functions	40.00%	35-40%	Solving for minimum
13	Geometry and Measurement	30.00%	25-30%	Consumer/ Producer Surplus
14	Geometry and Measurement	30.00%	25-30%	Demand, supply, etc.
15	Geometry and Measurement	30.00%	25-30%	Demand, supply, etc.
16	Geometry and Measurement	30.00%	25-30%	Consumer Surplus
17	Geometry and Measurement	30.00%	25-30%	Change in consumer surplus, producer surplus, or national welfare
18	Geometry and Measurement	30.00%	25-30%	Demand, supply, etc.
19	Data Analysis	10.00%	10-15%	Using circle graphs to interpret data
20	Data Analysis	10.00%	10-15%	Cost Tables

Table 2

	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
Non-consent	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)

Figure 1

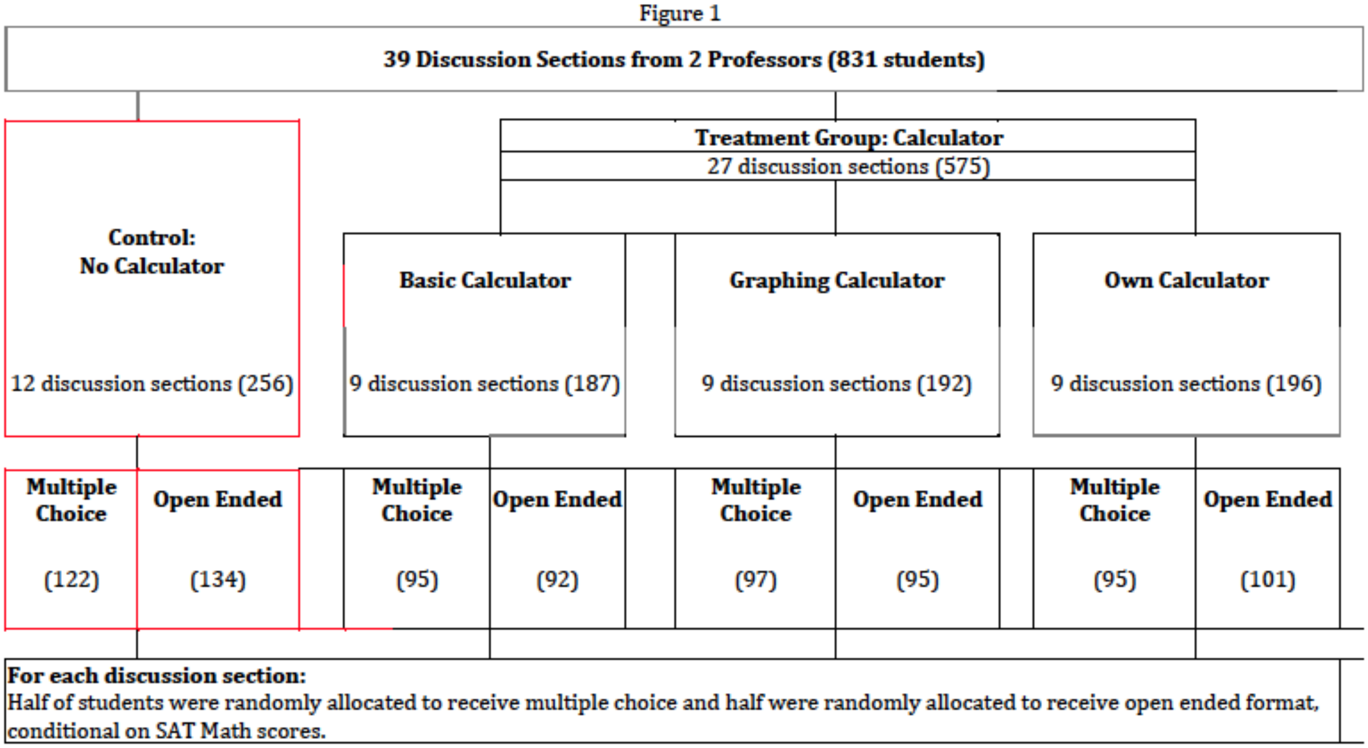
39 Discussion Sections from 2 Professors (831 students)
Treatment Group: Calculator
27 discussion sections (575)
Control:
No Calculator
12 discussion sections (256)
Basic Calculator
9 discussion sections (187)
Graphing Calculator
9 discussion sections (192)
Own Calculator
9 discussion sections (196)
Multiple Choice
(122)
Open Ended
(134)
Multiple Choice
(95)
Open Ended
(92)
Multiple Choice
(97)
Open Ended
(95)
Multiple Choice
(95)
Open Ended
(101)
For each discussion section:
Half of students were randomly allocated to receive multiple choice and half were randomly allocated to receive open ended format, conditional on SAT Math scores.

Table 3

	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

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

		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	δ1	.883***	
		(0.221)	
Basic Calculator	δ2	0.602***	236
		(.148)	(0.213)
Graphing Calculator	δ3	.471***	0.46***
		(0.159)	(0.137)
Own Calculator	δ4	.519***	0.218
		(0.154)	(0.152)
	α	.224*	1.1***
		(.112)	(0.128)
	N	544	412
	R^2	0.08	0.014

Table 6 Part I

		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 6 Part II

		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