

## The Crux: Promoting Success in Calculus II

### Abstract

In the 2013-14 school year, Boise State University (BSU) launched a major overhaul of Calculus I. The details of the reform, described elsewhere, involved both pedagogical and curricular changes. In subsequent years, we developed several assessment tools to measure the effects of the project on students' grades and retention. The toolkit includes: (1) pass rate and GPA in Calculus I, (2) longitudinal analysis of pass rates and GPA in subsequent courses, (3) impact of Calculus I on retention in STEM and retention at BSU, (4) all of the above comparing students in reformed Calculus vs traditional Calculus, (5) all of the above for underrepresented minorities, women, or other demographic subsets. While these tools were originally developed to study the Calculus I project, they are available for studying the effects of other courses on student academic performance and retention.

In this paper, we briefly describe a rebuild of Calculus II, overhauled in the 2015-16 school year following the same general plan as was used for Calculus I. We then present the results of applying the full toolkit to the new Calculus II course. Pass rate and GPA improvements in Calculus II were evident immediately after scale up in the spring of 2016. Sufficient time has now passed so that we can apply the full set of assessment tools built for Calculus I to measure the effectiveness of the Calculus II transformation on academic performance in post-requisite coursework and on student retention in STEM.

## 1.0 Introduction

The grade earned in mathematics courses is critical when considering student retention in engineering and in STEM majors. For example, the work by Budny, et al. (1998) shows that the grade earned in the first semester course of mathematics (whether Precalculus, Calculus I, or Calculus II) is a strong predictor of retention in engineering. Callahan & Belcheir (2017) showed that of the two – level of first semester mathematics, or grade earned – that the grade earned in the mathematics class is a better predictor of student retention in STEM one year later than the level of mathematics course taken. Success in the first year of mathematics in engineering is paramount.

Because of this, Boise State University is five years into an overhaul of the entry-year calculus sequence. Implementation of the initial, Calculus I, phase and early results were reported in Bullock, Callahan, Shadle (2015). This included pass rate gains that range from 8 to 10%, increased satisfaction by instructors, students and clients, and more. An examination of how students who have taken the overhauled Calculus I have fared in post-requisite coursework has been investigated in Bullock, Johnson, Callahan (2016) and Bullock, Callahan, Cullers (2017). The latter paper (2017) also presented the effects of the Calculus I project on retention.

As a natural next step in continuous improvement, the mathematics department turned to Calculus II as their next focal area for reform. In this paper, we report on what this reform of Calculus II consists of, and also track and report on student grade performance in the course as well as in post-requisite coursework including Dynamics, Fluids, Calculus III, and Differential Equations.

## 2.0 Background and Methods

### 2.1 Calculus II Redesign

The redesign of Calculus II followed the general plan that was used to redesign Calculus I (Bullock, et. al. 2015), with four major components of change:

1. Substantial changes to content, seeking to maximize relevance to future coursework.
2. Voluntary opt-in to a “master course” model.
3. Redesign of each daily lesson to support active learning pedagogies.
4. Formation of a community of practice to deliver the course.

The content of a typical second semester Calculus course usually includes: techniques of integration (symbolic with no machine assistance, plus some numerical integration), applications of integration (physical applications and solids of revolution), sequences and series (emphasizing proofs of convergence and culminating in Taylor series), and a smorgasbord of parametric functions, polar coordinates, conics, and differential equations. We rebuilt the content to focus tightly on four units:

- 4 weeks of symbolic integration. Restricted to a minimal list of types vetted by stakeholders.
- 4 weeks of sequences and series. No proofs. Qualitative understanding of convergence. Quantitative speed of convergence. Taylor polynomials as applied approximations.

- 4 weeks of applications of integration. Heavy emphasis on student understanding and communication of the underlying geometry (vs. formalism). Applications to loads, forces, moments, centroids, work, and energy.
- 3 weeks of 2-D parametric and vector valued functions. Mimicking the notation and language of the 3-D material that begins Calculus III at Boise State.

Previously, redesigned Calculus II was delivered as a collection of independent single sections with little to no governance beyond a common text and a suggestion of content coverage (the traditional list above). We replaced this with a master course specifying all homework, quizzes, exams, daily lesson order and content, and overall grade weighting. The master course was copied to each individual section, with the understanding that no instructor would be coerced by the department. Voluntary opt-in meant adopting the master course structure. We have had approximately 95% opt-in since the launch of the project.

Opt-in does not require any particular pedagogical approach. However, each homework set in the master course is designed to be best delivered in an active learning style, with most class time devoted to students progressing through carefully scaffolded exercises with guidance from the instructor and a learning assistant. All instructors who have opted in have also adopted some form of active learning.

The group of instructors in any given semester works as a team to deliver the course – collaborating on quizzes and exams, meeting regularly to discuss classroom practice and course delivery logistics. They are supported by a team of more senior instructors dedicated to the continued operation of the restructured Calculus I and II courses. The result is a strong community of practice.

Consensus and buy-in was developed over the 2015-16 scale up period by forming a Faculty Learning Community (FLC) that met for a full year (e.g. see Cox and Richlin, 2004). In the fall term of 2015, instructors debated and agreed upon lesson objectives and content. During the spring of 2016, all FLC members who had been assigned Calculus II taught their sections using the agreed upon curriculum and content. Weekly meetings during the spring semester served to further build out content, to discuss real-time issues in course delivery, and to agree on common weekly quizzes and midterm exams. These weekly meetings formed the basis for the ongoing community of practice that has continued the project. The result is a closely coordinated, multi-section Calculus II course with common content, assessments, and exams.

## 2.2 Methods

The toolkit developed to assess the effects of Calculus I transformation includes descriptive statistics:

- Time series of aggregate pass rate across all of Calculus I. (Bullock, et. al. 2015, 2016)
- Before/After comparisons of pass rates for individual instructors who taught both the old Calculus I and the reformed Calculus I. (Bullock, et. al. 2015)
- Pass rates in courses subsequent to Calculus I, with comparisons between students who reached the subsequent course via old Calculus I, reformed Calculus I, or by transfer credit. (Bullock, et. al. 2017)

- Retention of students, both in the sense of “retained at Boise State University” and “retained in STEM”, across the year in which they first encounter Calculus I, again with comparisons between old and reformed Calculus I. Effects on retention were studied for subpopulations of female, Pell eligible, and underrepresented minority students. Bullock, et. al. 2017).

Features of the course transformation process allowed us to identify a treatment group (those who took reformed Calculus I) and a control group (those who took the old Calculus I). The two groups co-existed across a time span that extended to either side of the year of course development and implementation. Before implementation, some students took the reformed curriculum as it was in preliminary development and testing, and after implementation, some instructors opted out of the project. The result is a natural experiment with two roughly equal sized study populations taking different versions of Calculus I in the same time frame. We used this opportunity to conduct the following statistically rigorous assessments:

- Comparison of Calculus I pass rates for treatment vs. control. Significance tests were applied to the research question: “Does treatment improve Calculus I pass rate?” Control variables were used to test whether the two groups had different levels of academic preparation or ability. (Bullock, et. al. 2016)
- Comparison of Calculus I average GPA for treatment vs. control. Significance tests were applied to the research question: “Does treatment improve Calculus I GPA?” Control variables were used to test whether the two groups had different levels of academic preparation or ability. (Bullock, et. al. 2016)
- Comparison of Calculus II pass rates and GPA for treatment vs. control. Note that Calculus II in this context is not the treatment course. It is a testing ground for the results of reforming Calculus I. Significance tests were applied to the research question: “Does the treatment (reformed Calculus I) have any detrimental effects on Calculus II?” Control variables were used to test whether the two groups had different levels of academic preparation or ability. (Bullock, et. al. 2016)
- Comparison of some (not all) of the various retention metrics. Significance tests were applied to the research question: “Does treatment improve retention?” Control variables were not used, so this is less rigorous. The significance testing here is perhaps best thought of as a refinement of the descriptive statistics on retention. (Bullock, et. al. 2017)

In the subsequent sections of this paper we will, for each assessment instrument or group above, present the results of applying the same tools or tests to measure the effects of transforming Calculus II. In each case, we will compare or contrast the findings with what we learned about Calculus I across the last three years.

### **3.0 Results – Descriptive Statistics**

#### **3.1 Aggregate Pass Rate**

Figure 1 shows the pass rate for all of Calculus II in each non-summer term for the last decade (line graph). The bars graph shows total enrollment. Color coding indicates students in old Calculus II (blue) versus new Calculus II (orange). The implementation term is visible in the shift from mostly blue to mostly orange bars. Orange before implementation is due to small development and testing sections. Blue after implementation is due to instructors opting out of

the coordinated course design. Despite some volatility and a potential trend leading up to transition, there is a fairly clear jump in pass rate.

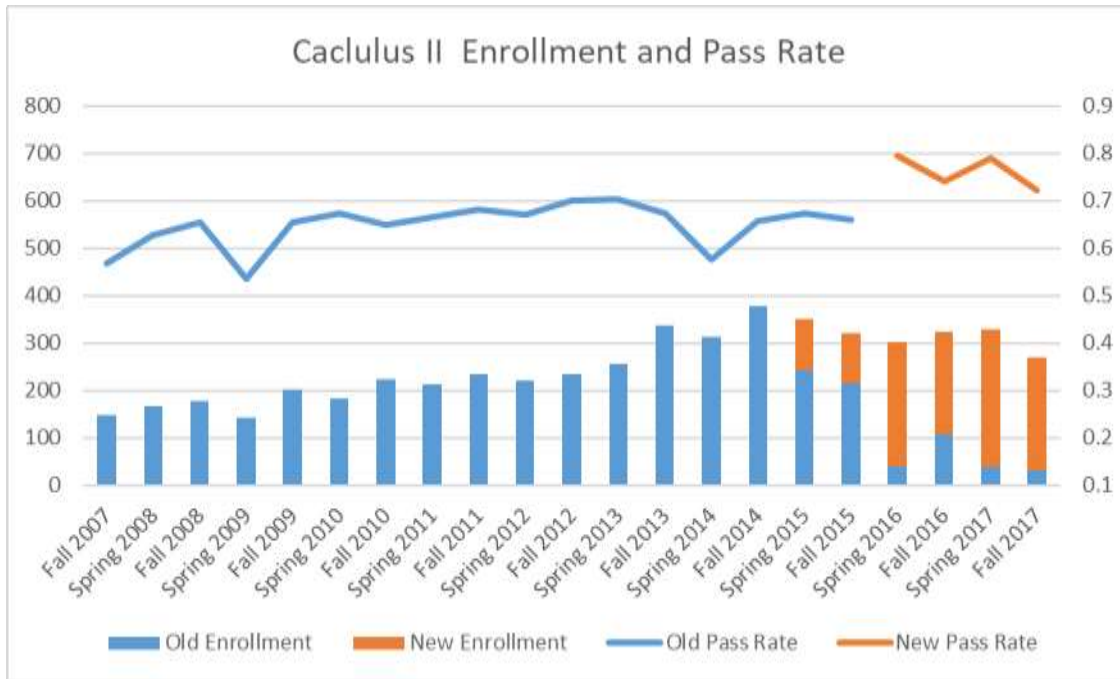


Figure 1: Calculus II enrollment and pass rate

For comparison, Figure 2 shows the corresponding decade of Calculus I.

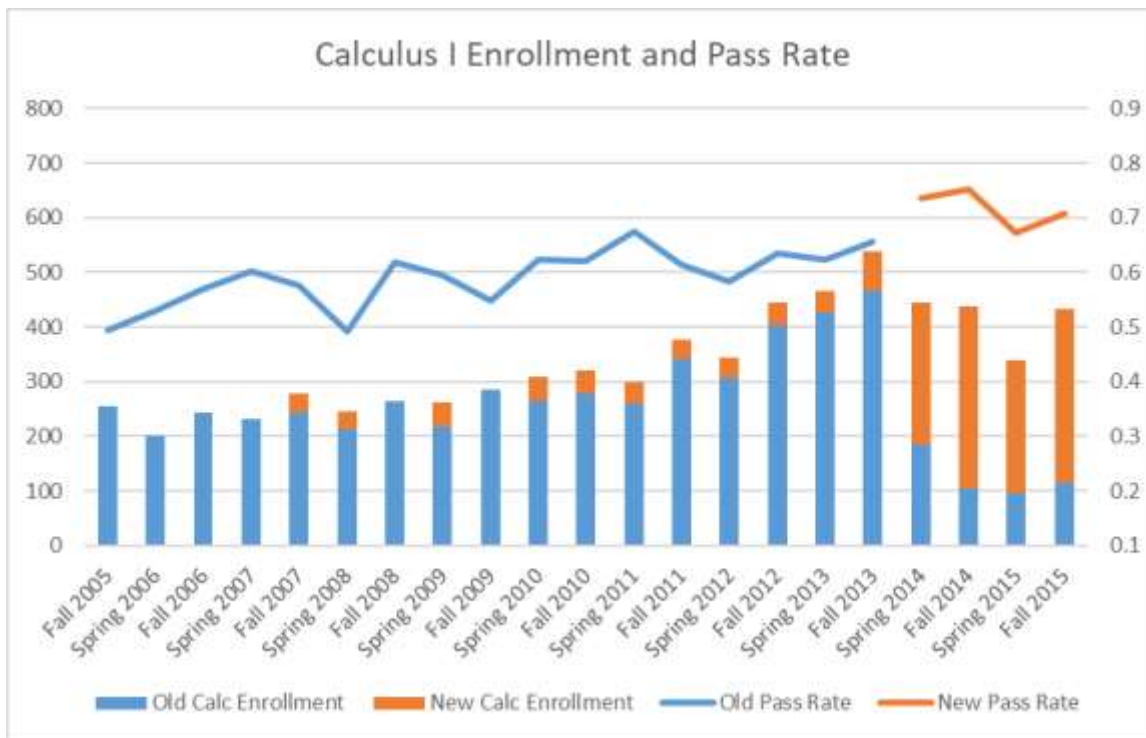


Figure 2: Calculus I enrollment and pass rate

In both graphs we have chosen to present one decade of data, with the cut off exactly 2 years after the course was transformed. Both graphs show that after transition, the bulk of calculus was taught using the reformed curriculum, and pass rates increased.

### 3.2 Before and After Pass Rate

In the Calculus I transition, fortuitously, we had a group of six instructors who taught Calculus I both before and after the reform, allowing us to compare pass rates while keeping instructors constant. In the Calculus II transition we, coincidentally, ended up with six instructors who had taught both the old and new Calculus II. Figures 3 and 4 show the individual pass rates, per instructor, for both Calculus I (Bullock, et. al., 2015) and Calculus II.

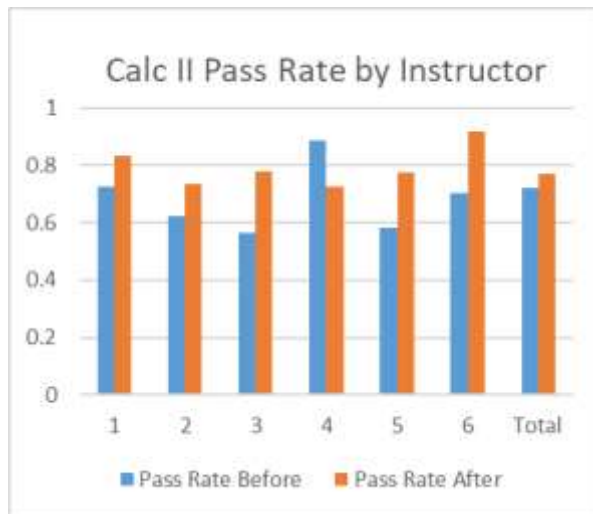
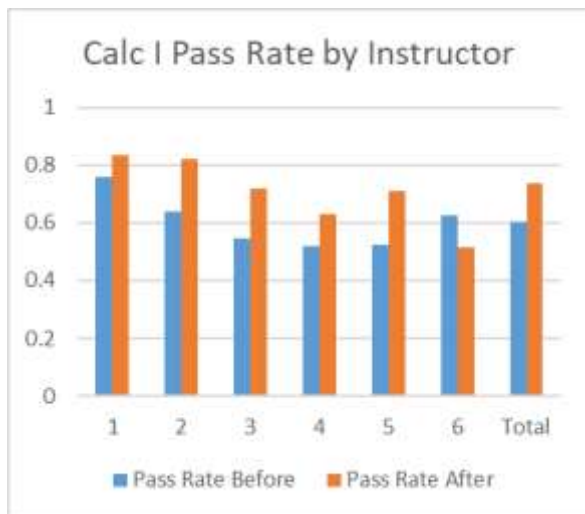


Figure 3: Calculus I pass rate by instructor      Figure 4: Calculus II pass rate by instructor

For both Calculus I and II, five of the six instructors saw jumps in pass rate. However, this data is highly volatile, with small population sizes. The rightmost bar aggregates the pass rate across the six instructors, giving a decent comparison of before/after pass rates while holding the instructor corps constant.

### 3.3 Subsequent Course Work

As an assessment of the efficacy of Calculus II, we monitor pass rates in courses that carry Calculus II as a prerequisite or for which Calculus II knowledge could be meaningful even if not a prerequisite. We consider all students who took and passed Calculus II between Spring 2015 and Summer 2017. This range is chosen to include a full calendar year before the implementation term (Spring 2016) for transforming Calculus II, and to end at the last point when a student could pass Calculus II and subsequently attempt another course. In this time frame, there is one data record for each pair of events of the form:

(Student Passed Calculus II, Same student subsequently attempted a target course)

A student can appear more than once in the data set, if they have attempted more than one of the subsequent target courses. All students attempting any given target course are sorted by whether they passed new Calculus II or old Calculus II. We compute the pass rate for each group in each

course. For comparison, we also include the pass rate for students who transferred the prerequisite. These students have no record of a Boise State Calculus II course prior to the target course, so they are not affected by our course redesign. Results are shown in Table 1.

Table 1: Post Calculus II pass rates -- individual courses

Post Calculus II Pass Rates -- Individual Courses						
Course	Old Calculus II	New Calculus II	Transfer	Effect Size	$p$ -value	$N$
Calculus III	84.2%	84.2%	80.6%	0.0%	0.988	811
Circuits I	80.0%	81.3%	71.4%	1.3%	0.876	98
Circuits II	65.0%	83.9%	100.0%	18.9%	0.133	51
Diff Eq	80.9%	76.0%	70.7%	-4.9%	0.154	582
Dynamics	84.0%	77.7%	84.2%	-6.3%	0.240	212
E and M	72.7%	75.0%	100.0%	2.3%	0.901	23
Fluids	84.4%	89.5%	61.9%	5.1%	0.403	121
Heat	92.9%	88.2%	50.0%	-4.6%	0.616	45
Mech Mat	81.2%	77.2%	87.0%	-4.0%	0.532	164
Phys I	89.0%	89.7%	88.7%	0.8%	0.850	235
Phys II	88.7%	91.5%	81.3%	2.8%	0.278	519
Statics	75.8%	78.9%	64.4%	3.1%	0.477	360
ALL COURSES	83.3%	82.8%	77.9%	-0.5%	0.728	3221

For those who prefer a graphical description, see Figure 5.

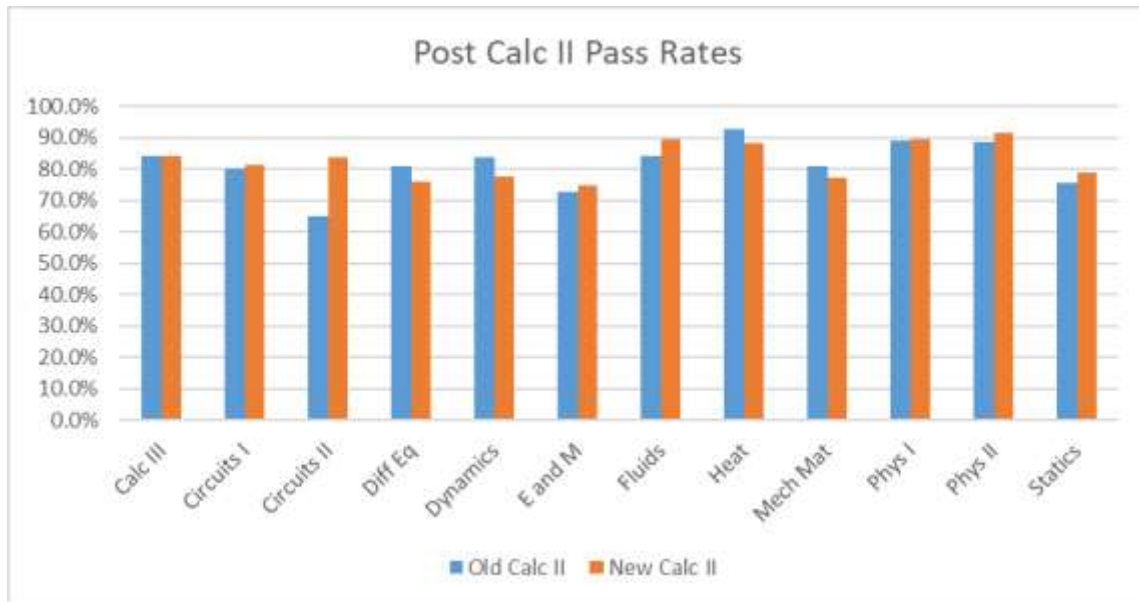


Figure 5: Post Calculus II pass rates

In Table 1, effect size is the difference between the pass rates of students originating in new Calculus II compared to old Calculus II. Positive effects mean that the new Calculus II students perform better.  $N$  and  $p$ -value are included to help judge significance. However, since none of the effects are significant, this is simply additional descriptive statistics. For example, if there is a negative effect with a small  $p$ -value, even if not meeting the 0.05 significance threshold, this is a potential cause for concern. Details on the computational methods are available in Bullock, et. al. (2017).

The purpose of this computation is to give a sense of whether the changes to curriculum and content in the new Calculus are creating any problems in downstream courses. Since the content changes have made Calculus II more accessible, there is some possibility that subsequent coursework would expose students' weaknesses. Since we see a scattering of positive and negative effects, but none statistically significant, this descriptive report suggests that there are no ill effects.

This tool allows for easy aggregation of post Calculus II courses by discipline, which is of interest to specific course owners. It also includes demographic slicers. The discipline aggregates are Math, Physics, and Engineering (Table 2).

*Table 2: Post Calculus II pass rates -- by discipline*

<b>Post Calculus II Pass Rates – By Discipline</b>						
Discipline	Old Calculus II	New Calculus II	Transfer	Effect Size	$p$ -value	$N$
Engineering	79.9%	80.3%	73.8%	0.1%	0.952	1189
Math	82.9%	80.6%	75.3%	-2.7%	0.195	1530
Physics	88.8%	90.9%	85.7%	2.0%	0.373	866

The subpopulations of most interest to us are women, underrepresented minorities (URM), and Pell eligible students. For this, we aggregate post Calculus II courses (Table 3).

*Table 3: Post Calculus II pass rates -- by demographic*

<b>Post Calculus II Pass Rates -- By Demographic</b>						
Demographic	Old Calculus II	New Calculus II	Transfer	Effect Size	$p$ -value	$N$
URM	79.8%	80.4%	66.7%	0.6%	0.879	410
Female	88.7%	88.1%	84.7%	-0.6%	0.803	699
Pell	81.3%	78.3%	80.5%	-3.0%	0.254	915

As always, these are descriptive statistics, with  $N$  and  $p$ -value included to provide suggestions of which numbers might be of most interest.



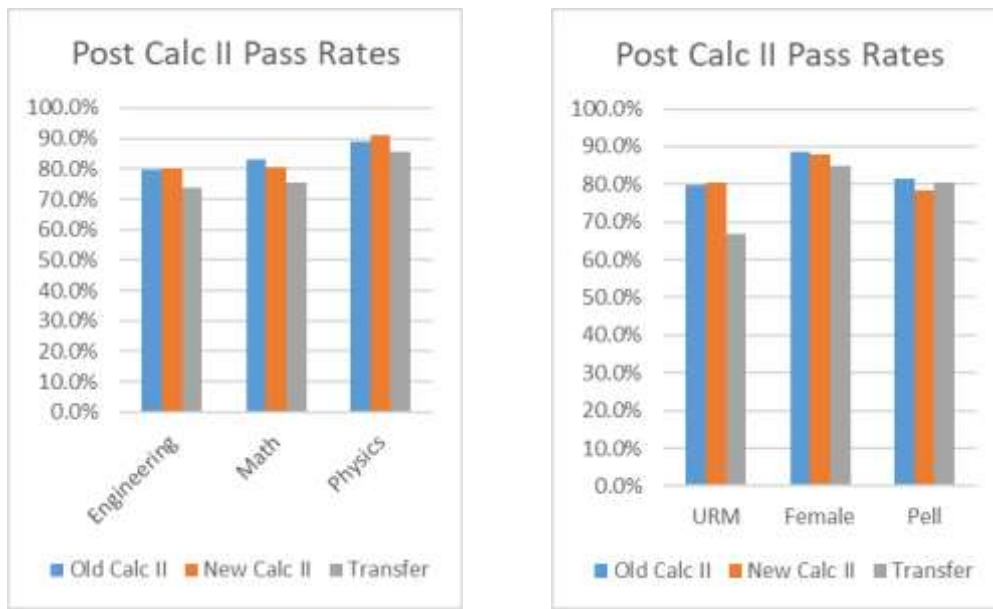


Figure 6: Post Calculus II pass rates - by discipline and by demographic

## 4.0 Results – Rigorous Hypothesis Tests

### 4.1 Outcomes in the Transformed Course

Our first use of statistical testing of hypotheses to address a research question in the Calculus I project was a comparison of outcomes in Calculus I for treatment (new Calculus I) vs control (old Calculus I). Details of the methodology are in Bullock, et. al. (2016), where we found large and significant gains in pass rate and GPA in Calculus I. For this paper, we applied the same methodology to treatment and control populations of Calculus II students. The study population was all students in Calculus II from Spring 2013 through Fall 2017, a four-year span straddling the implementation term, Spring 2016. There were 2845 data records, split into 1307 treated students and 1538 in the control group. The research question was:

“Does treatment (reformed Calculus II) improve results in Calculus II?”

We tested two null hypotheses. Regarding pass rates:

*H<sub>0</sub>: Students in treatment and control are equally likely to pass Calculus II.*

Regarding grades:

*H<sub>0</sub>: Treatment and control groups will have the same average grades in Calculus II.*

The experimental variables we measured were Pass Rate and Average Grade Points (GPA) for each group in Calculus II. We also sought to control for the possibility that the treatment and control groups had different levels of academic preparation or aptitude. For each group, we measured four additional variables: High School GPA, College GPA (in the term they took Calculus II), Admission Index (computed by our admissions office from HS GPA and composite SAT and/or ACT scores), and ACT Math score, using concordances if a student has an SAT Math score instead. The results are shown in Table 4.

Table 4: Calculus II pass rate and GPA, treatment vs. control

Calculus II Pass Rate and GPA: Treatment vs. Control					
Variable	Variable Type	Control	Treatment	Effect Size	<i>p</i> -value
Calculus II Pass Rate	Study Variable	63.7%	77.4%	13.6%	0.0000
Calculus II GPA	Study Variable	1.90	2.38	0.48	0.0000
College GPA	Control Variable	3.01	3.09	0.08	0.0002
Admission Index	Control Variable	62.42	63.98	1.56	0.0708
High School GPA	Control Variable	3.35	3.40	0.06	0.0061
Concordant ACT Math	Control Variable	25.14	25.59	0.45	0.0160

It is immediately evident that there are massive gains in pass rate (13.6%) and GPA (an increase of half a letter grade) for the treatment group. However, it is also clear that the treatment group in this natural experiment is stronger in academic preparation. We have used this “academic preparation control” process in all of the previous Calculus I papers – and in each case, we found that treatment and control groups were not academically different, so we were satisfied with this form of control. However, the results in Calculus II make it clear that better tools are needed – either a multivariable regression to determine what portion of the gains are due to treatment instead of incoming academic ability, or perhaps non-parametric methods. Unfortunately, this will have to wait for a subsequent study. For now, we can report enormous gains with statistical significance on the study variables. These are more than twice as large as the gains shown in Calculus I at the equivalent stage of that project. If even half of the Calculus II gains are due to the treatment, this is still an excellent outcome.

#### 4.2 Outcomes in Subsequent Courses

Section 3.3 provided descriptive statistics on pass rates in courses subsequent to the transformed Calculus II course. We can also use the tool to address the research question:

“Does treatment (reformed Calculus II) have any negative effect on subsequent courses?”

Essentially, this is a test of “do no harm.” Early in the Calculus I project, there was some fear that pass rate gains in Calculus I might be coming at the expense of success in subsequent courses, so we built and applied this tool as a rigorous test to check if there was any harm. We found none for the Calculus I reform. Similarly, for Calculus II, we test the null hypothesis:

*H<sub>0</sub>: Treatment and control groups (in Calculus II) are equally likely to pass subsequent courses.*

Here, we hope to find no evidence that causes us to reject the null hypothesis. We set up a natural experiment in Calculus II following exactly the protocol we used for Calculus I (Bullock, et. al. 2016). In that paper, we tested only the pass rate in one critical course subsequent to Calculus I – namely Calculus II. However, with Calculus II as the treatment focus, there is less clarity as to what subsequent course is the most important test of treatment effects. We chose two: Calculus III and Differential Equations. Both courses are part of the standard STEM track; either course may be taken immediately after Calculus II. Which comes first is typically a matter of advising within various STEM disciplines. There are additional technical details of how we restricted the study population to most effectively test our hypothesis, which we will not restate here (see Bullock, et. al. 2016). Tables 5 and 6 present the results for the two subsequent courses.

Table 5: Post Calculus II results in Calculus III

Post Calculus II results in Calculus III					
Variable	Variable Type	Control	Treatment	Effect Size	<i>p</i> -value
Calculus III Pass Rate	Study Variable	84.3%	83.8%	-0.4%	0.884
Calculus III GPA	Study Variable	2.64	2.59	-0.05	0.620
College GPA	Control Variable	3.19	3.19	0.00	0.937
Admission Index	Control Variable	66.95	67.34	0.39	0.802
High School GPA	Control Variable	3.48	3.47	-0.01	0.781
Concordant ACT Math	Control Variable	25.45	25.89	0.43	0.236

In Calculus III, both study variables show a small negative effect of treatment, but very large *p*-values mean this is insignificant, so the null hypothesis of “did no harm” is retained. This is what we found when we studied the effect of Calculus I on subsequent Calculus II. Also, similarly, the treatment and control groups display no significant differences in academic ability or preparation.

The picture for Differential Equations, however, is less appealing.

Table 6: Post Calculus II results in Differential Equations

Post Calculus II results in Differential Equations					
Variable	Variable Type	Control	Treatment	Effect Size	<i>p</i> -value
Diff Eq Pass Rate	Study Variable	80.3%	70.4%	-9.9%	0.068
Diff Eq GPA	Study Variable	2.46	2.16	-0.30	0.078
College GPA	Control Variable	3.06	3.11	0.06	0.378
Admission Index	Control Variable	63.27	64.00	0.73	0.796
High School GPA	Control Variable	3.38	3.41	0.03	0.638
Concordant ACT Math	Control Variable	24.67	26.53	1.86	0.008

Here, we see very large negative effects on the treatment population. While the *p*-value is just above the threshold at which one would typically reject the null hypothesis, it would not be safe to comfortably conclude that the treatment of reforming Calculus II has done no harm in Differential Equations. Also, since there is evidence in the control variables that indicates the treatment group was academically stronger than the control group, it puts the negative treatment effects in an even worse light. Again, it is clear that a more robust statistical model is necessary. But this data is sufficient to require immediate engagement with the Calculus II project team and possible intervention to ameliorate potential trouble in Differential Equations. It is unclear what causal mechanism (if any) may be at work.

### 4.3 Retention

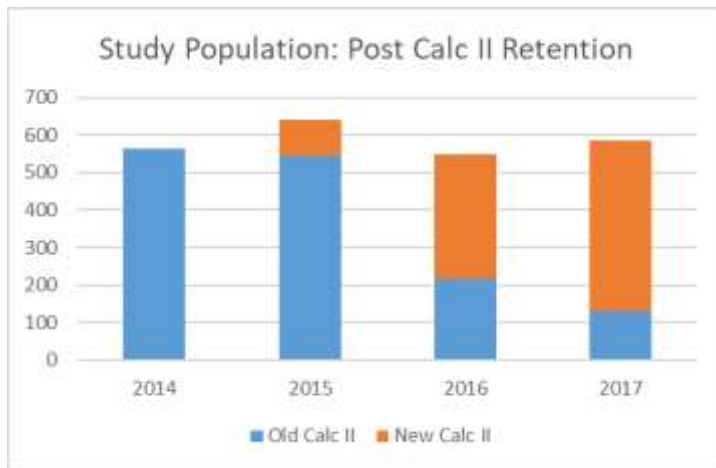
Here again we developed a natural experiment as the Calculus I project evolved (Bullock, et. al. 2017). We used the experiment to study the effect of reforming Calculus I on the retention of students in the year that they encountered Calculus I. For this paper, we apply an identical protocol to Calculus II students, addressing the research question:

“What effect does treatment (reformed Calculus II) have on retention of students in the year that they encounter Calculus II?”

There are actually two research questions: one in which retention is “retained at the university,” regardless of major, and one in which retention is “retained in STEM,” and applies only to students who were STEM majors in the year they encountered Calculus II<sup>1</sup>. We answer each question for the general study population and then again for demographics of women, underrepresented minorities, and Pell eligible students. In all cases, we test the null hypothesis:

*H<sub>0</sub>: Students in treatment and control are equally likely to be retained.*

We do not, however, include the additional variables for academic preparation and ability.



Details on the protocol for forming the study population, technical definitions of variables, and other elements of the experimental design can be found in Bullock, et. al. (2017). Figure 7 presents a snapshot of the size of the study population (2340 records), distributed across 4 academic years and broken out as treatment (new Calculus II) or control (old Calculus II).

Figure 7: Study population – post Calculus II retention

### 4.3.1 Retained at the University

Treatment delivers a bit more than four percentage points of additional retention at the university in the year that students encounter Calculus II (Table 7 and Figure 8). The result is statistically significant.

Table 7: Post Calculus II retention rates

Post Calculus II Retention Rates					
Demographic	Control	Treatment	Effect Size	p-value	N
ALL	81.0%	85.2%	4.2%	0.008	2340
Female	83.1%	88.2%	5.1%	0.104	496
URM	81.5%	85.3%	3.7%	0.372	324
Pell	82.6%	84.9%	2.3%	0.400	789

When sliced by demographics, we see that there are slightly larger retention gains for women. URM and Pell eligible students also gain, but not as much as the full study population. None of

<sup>1</sup> Our definitions of the terms “retention” and “retention rate” differ from the definitions used in Boise State University’s official reporting offices. Details available in Bullock, et. al. (2017).

the demographically specific gains are statistically significant, since these are much smaller populations compared to the full study population.

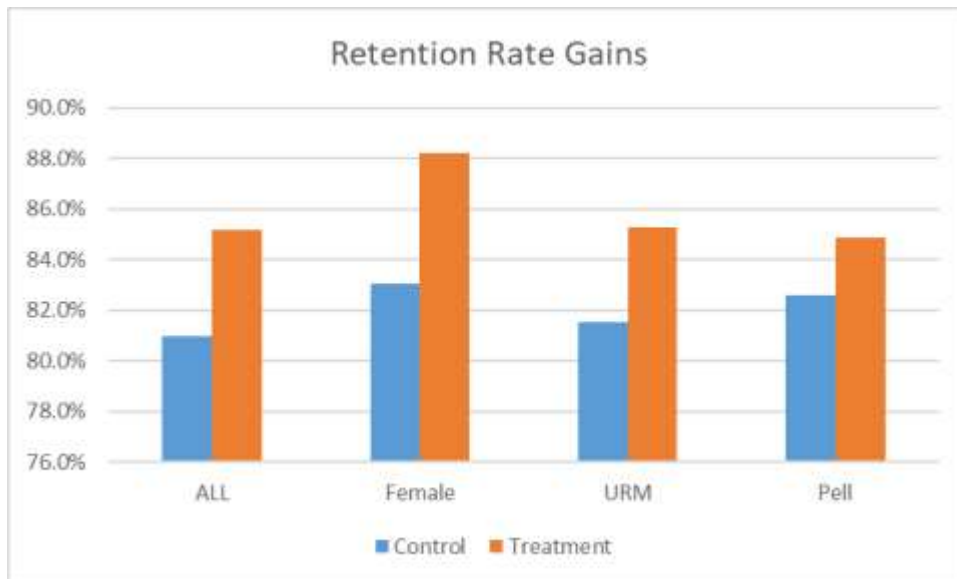


Figure 8: Retention rate gains, Calculus II

### 4.3.2 Retained in STEM

We restrict the study population to students who were STEM majors in the year they encountered Calculus II. There are now three possible outcomes: Retained in STEM, switched to non-STEM, and left school. Treatment delivers a similar gain in STEM-to-STEM retention.

Table 8: Post Calculus II STEM Retention

Post Calculus II STEM Retention					
Result	Control	Treatment	Effect Size	<i>p</i> -value	<i>N</i>
STEM-to-STEM	75.5%	79.8%	4.3%	0.040	1659
STEM-to-Non	5.9%	6.0%	0.1%		128
Dropped Out	18.5%	14.2%	-4.3%		363

This is very much like what we saw for Calculus I (Bullock, et. al. 2017) in two ways. One is that the size of the gain is what one would expect as a simple consequence of the pass rate gains, and two is that the entire gain in STEM-to-STEM retention is caused by preventing dropouts. Both observations suggest that all of this is directly attributable to pass rate.

When we drill down to demographics (Table 9) we see similar results, albeit none that are statistically significant. There is one notable difference involving underrepresented minority students.

For female students, treatment may confer a gain in STEM retention that is, again, entirely the result of preventing dropouts. The STEM-to-STEM retention gain for women is not as large as the gain in retention at the university, which is a stark contrast to the result from transforming Calculus I (Bullock, et. al. 2017). In that paper we found a much larger benefit to women retained in STEM as compared to women retained at the university. Also, note that the starting

point for female retention in STEM, about 75%, is much lower than the starting point for female retention in college

Table 9: STEM retention by Demographic

STEM Retention by Demographic						
Demographic	Result	Control	Treatment	Effect Size	p-value	N
Female	STEM-to-STEM	74.8%	78.5%	3.7%	0.435	328
	STEM-to-Non	8.9%	9.3%	0.4%		39
	Dropped Out	16.3%	12.2%	-4.1%		63
URM	STEM-to-STEM	77.2%	76.9%	-0.3%	0.953	225
	STEM-to-Non	5.3%	9.9%	4.7%		21
	Dropped Out	17.5%	13.2%	-4.3%		46
Pell Eligible	STEM-to-STEM	76.6%	79.5%	3.0%	0.403	576
	STEM-to-Non	6.1%	6.1%	0.0%		45
	Dropped Out	17.4%	14.4%	-3.0%		121

For Pell eligible students, there is the same story: small gains that are due to preventing dropouts.

There is an oddity for URM. Here, the treatment effect on STEM retention is negative. Reforming Calculus II could have cost some URM retention in STEM. As with other groups, we have obtained a nice reduction in the dropout rate, but here all of the non-dropouts seem to have departed for non-STEM fields.

While informative, none of the demographically specific results are statistically significant.

### 4.3.3 STEM Retention Gaps

The previous subsection details STEM retention rates for demographic subgroups, which can be compared to STEM retention for the full study population.

Table 10: STEM retention gaps

STEM Retention Gaps		
Demographic	Control	Treatment
Female	74.8%	78.5%
Male	75.7%	80.2%
<b>Gender Gap</b>	<b>0.9%</b>	<b>1.7%</b>
URM	77.2%	76.9%
non-URM	77.8%	82.0%
<b>URM Gap</b>	<b>0.6%</b>	<b>5.1%</b>
Pell	76.6%	79.5%
non-Pell	78.5%	82.0%
<b>Pell Gap</b>	<b>2.0%</b>	<b>2.5%</b>

Where retention gaps are concerned, what is more appropriate is a head-to-head comparison. These are displayed for treatment and control in Table 10. Here, we show only STEM-to-STEM retention. It is evident that the treatment seems to confer STEM retention gains for all groups. However, because the gains for men are highest, the pre-existing gaps for women, underrepresented minorities, and Pell eligible students widened after treatment.

## 5.0 Summary

The transformation of Calculus II has achieved very large gains in Calculus II pass rates and grades, which translate into reasonably large gains in retention, both at the overall university level and specifically for STEM majors. All of these results, when studied via natural experiment, are statistically significant. None are restricted to *a priori* advantaged demographic groups. The gains in pass rate, grades, and retention are similar to those achieved by the earlier transformation of Calculus I at Boise State University. The Calculus II gains are even larger.

Descriptive statistics on performance in courses beyond Calculus II suggest that there is no negative effect from altering the Calculus II content and curriculum. However, when statistical tools are carefully applied to test this hypothesis on immediately subsequent math courses, there is one important and actionable exception; although Calculus II transformation seems to have no effect on Calculus III, there is a sizable and significant negative impact on Differential Equations. It is, at least, a positive outcome of this study to have caught this effect and to have data to support and guide an intervention to address it.

Retention effects are smaller and less statistically robust than the pass rate gains in Calculus II. They also did not display STEM specific impacts that were as profound as those observed after Calculus I transformation. However, this does not mean that the Calculus II reform is failing female, URM, or Pell eligible students. It simply means that issues with retention will need to be kept in view.

## 6.0 Acknowledgments

The authors gratefully acknowledge the support of the National Science Foundation through Grant No DUE-1347830, the ongoing support of the Dean of Arts & Sciences and the Office of the Provost at Boise State University, and the reviewers for suggestions that improved the paper.

## References

- Budny, D., LeBold, W., Bjedov, G. (1998). Assessment of the Impact of Freshmen Engineering Courses. *Journal of Engineering Education*, 87(4) 405-411.
- Bullock, D., Callahan, J., Shadle, S. E. (2015). Coherent Calculus Course Design: Creating Faculty Buy-in for Student Success. *Proceedings of the 2015 ASEE Annual Conference & Exposition*, Seattle, WA.
- Bullock, D., Johnson, K. E., Callahan, J. (2016). Longitudinal Success of Calculus I Reform. *Proceedings of the 2016 ASEE Annual Conference & Exposition*, New Orleans, LA.
- Bullock, D., Callahan, J., Cullers, J. B. S. (2017). Calculus Reform - Increasing STEM Retention and Post-Requisite Course Success While Closing the Retention Gap for Women and Underrepresented Minority Students. *Proceedings of the 2017 ASEE Annual Conference & Exposition*, Columbus, OH.
- Callahan, J., Belcheir, M. (2017). Testing our Assumptions: The Role of First Course Grade and Course Level in Math and English in Predicting Retention. *Journal for College Student Retention*, 19(1).
- Cox, M.D., Richlin, L. (Eds.) (2004). *New directions for teaching and learning, No. 97. Building faculty learning communities*. San Francisco, CA, Jossey-Bass.