Math Pre-requisite Study: First Stage

Jeff Webb, SLCC IR

August 31, 2015

1 Introduction

The math department has pre-requisites for (nearly all) math courses. Pre-requisites are supposed to expire after a year but the department has no way of enforcing these pre-requisites, with the consequence that some students with expired pre-requisites inappropriately enroll in courses. Math has done some analysis to suggest that as the time since pre-requisites increases, student performance declines. The department would like this analysis confirmed. Specifically, the research questions include:

1. Does the length of time since completing a pre-requisite impact student performance (measured in terms of passing grade or course completion)?
2. If there is an impact, what is the ideal cut off point for expiration to ensure student performance while reasonably accommodating gaps in the sequence. Currently the expiration is set at 2 semesters.

The purpose of this draft is to:

- Describe modelling choices we've made at the level of the data.
- Present visualizations and tables summarizing course performance and course completion in relation to the time between math courses.
- Estimate the effect size of enforcing the expiration on pre-requisites. Suppose that enforcement makes a difference to pass/completion rates—is that difference likely to be practically meaningful?

Any statistical modelling we decide to do—based on conversations arising from the results here—will be presented in subsequent work.

2 Data

The dataset includes all students who took a math course from Fall 2009 through Summer 2015. The variable of interest for this analysis was the gap between math classes taken by a student. Rather than tracking semesters since pre-requisites, we were interested only in the gaps between math classes. The reasoning here was that even if the last class did not happen to be a pre-requisite, it would nevertheless serve the purpose of a pre-requisite, if passed, by keeping the student engaged in math. This modeling choice reflects the math department's practice of waiving the pre-requisite in such cases. There were additional modeling choices:

- Time between math classes was measured in terms of semesters. This "time between" or gap variable counts the semesters between classes. The gap would be 0 in the case of adjacent semesters.
- A student's last class was used for calculating the gap variable only if the grade was C or better.
- Students who audited were removed from the dataset.
- Students who withdrew (grade of W) were counted as neither completing nor passing their courses.

- Non-completion was defined by a grade of W, E or I.
- If a student took two math courses simultaneously, both courses received the same gap count derived from the last passed math class.

Although the data starts with Fall 2009, if a student took a math class before that date, the earlier date was used to calculate the gap variable.

2.1 Examples of student records

id	term	course	grade	time_between
2288	201340	1010	D	NA
2288	201420	1010	C+	NA
2288	201430	1050	W	0
2288	201430	1030	B	0

id	term	course	grade	time_between
17976	201040	950	B-	NA
17976	201120	990	W	0
17976	201320	995	P	6
17976	201320	990	B+	6
17976	201340	1010	D	1
17976	201420	1010	B-	2
17976	201440	1050	E	1
17976	201520	1050	W	2

id	term	course	grade	time_between
5432	201120	920	A	NA
5432	201230	990	A	3
5432	201240	1010	E	0

id	term	course	grade	time_between
565348	200940	920	B	NA
565348	201020	990	C+	0
565348	201030	1010	C+	0
565348	201120	1050	D-	1
565348	201140	1050	B	3
565348	201220	1060	W	0
565348	201240	1060	D	2
565348	201320	1060	E	3
565348	201340	1060	C-	5
565348	201420	1060	E	6
565348	201430	1060	D	7
565348	201440	1040	W	8
565348	201440	1060	E	8
565348	201520	1060	D	9
565348	201520	1040	D	9

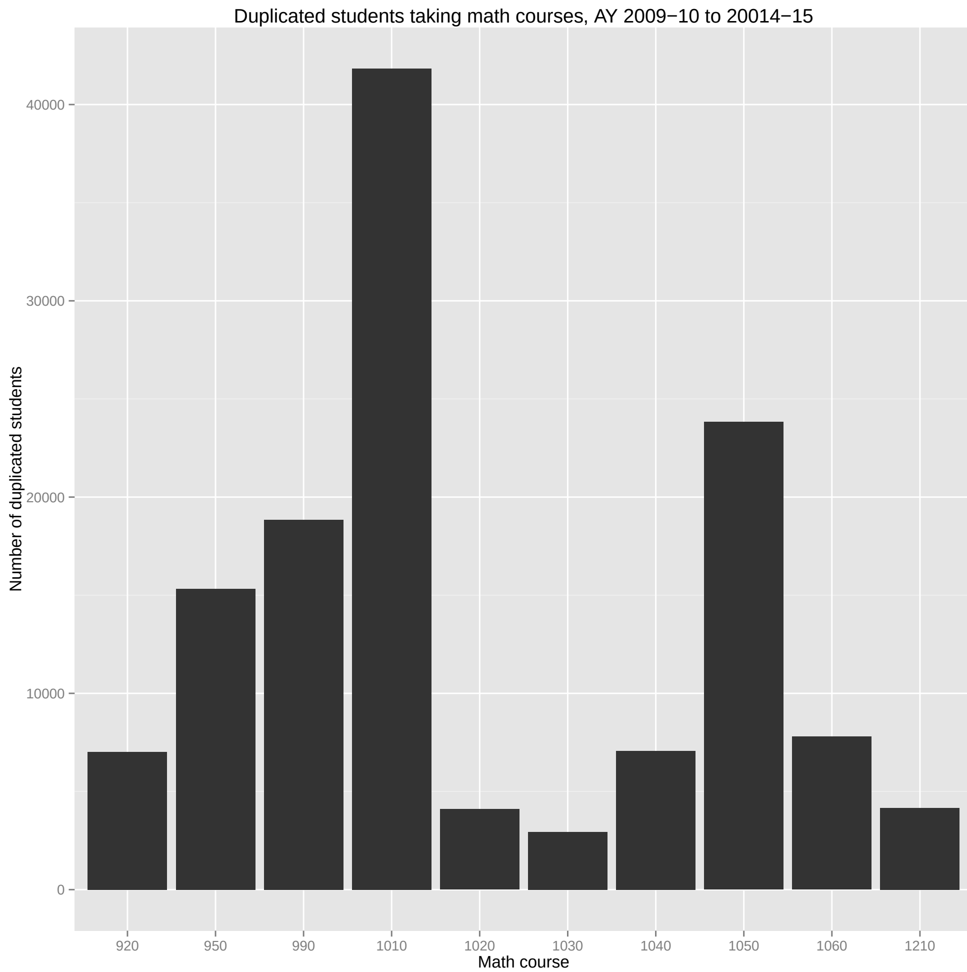
id	term	course	grade	time_between
14237	201120	1010	C+	9
14237	201140	1050	D	1
14237	201220	1050	B	2
14237	201240	1060	E	1
14237	201320	1060	E	2
14237	201330	1060	E	3

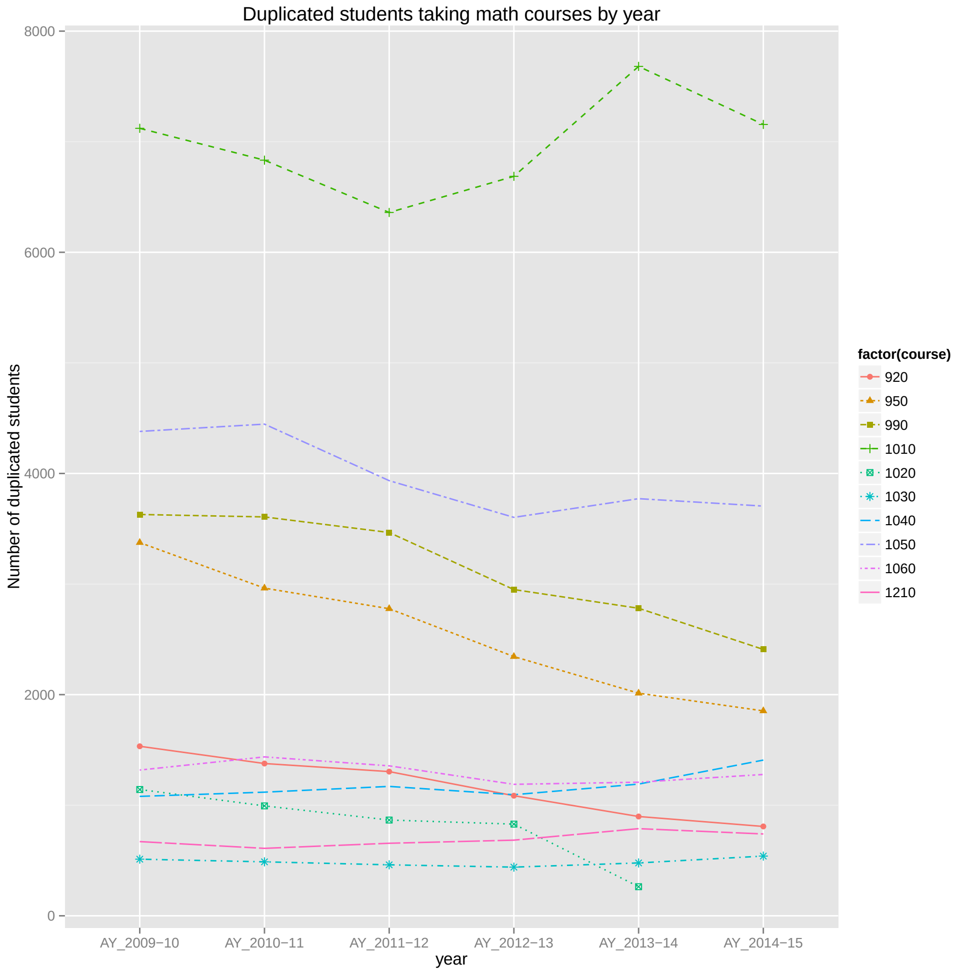
id	term	course	grade	time_between
565898	200940	920	E	NA
565898	201130	925	P	NA
565898	201130	920	A-	NA
565898	201140	995	P	0
565898	201140	990	C	0
565898	201220	1010	E	0
565898	201240	1010	B	2
565898	201240	1015	P	2
565898	201320	1055	P	0
565898	201320	1050	E	0
565898	201330	1050	C	0
565898	201340	1050	C	0
565898	201420	1060	E	0
565898	201430	1060	D+	1
565898	201440	1060	C+	2

Note: The first gap variable in each record is NA unless there was an enrollment that preceded Fall 2009.

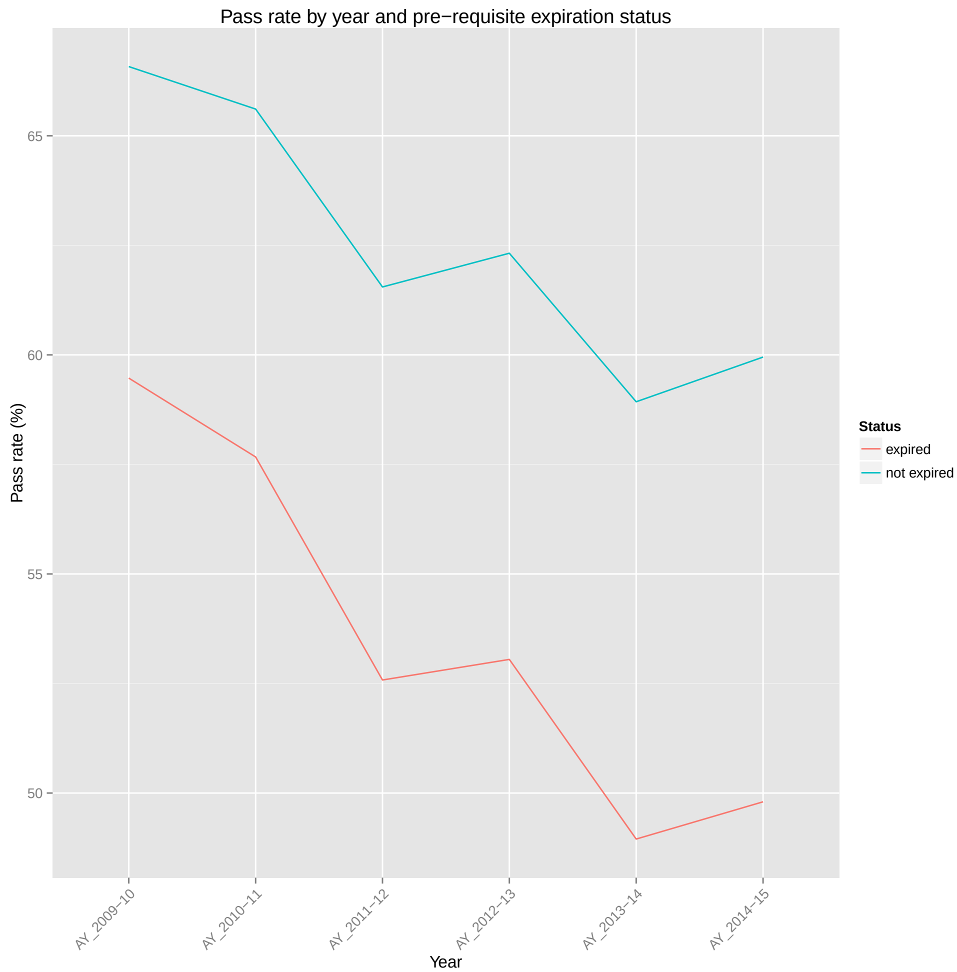
3 Descriptive statistics

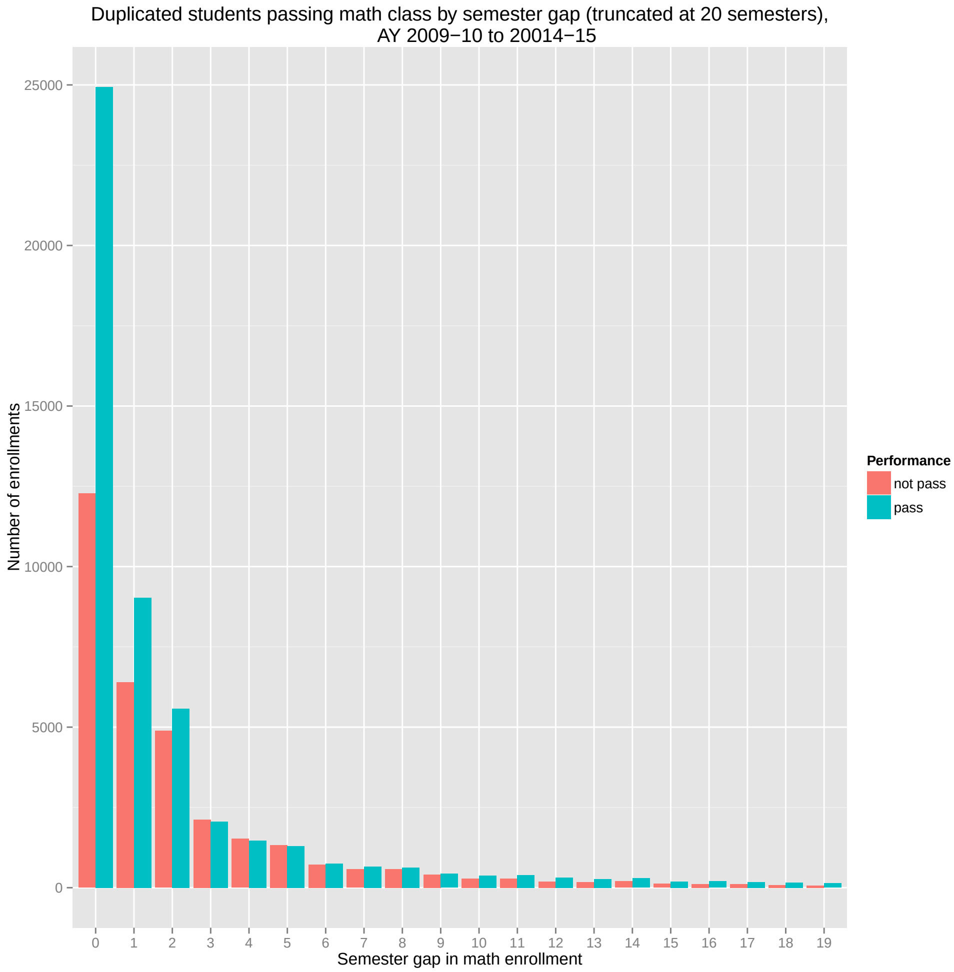
In this time period 5123 courses were offered with 145988 enrollments for 64307 unique students. The average number of math classes taken by students was 2.27. Here are plots of duplicated students taking math courses from academic year 2009-10 to 2014-15 inclusive, aggregating over years and by year.


Duplicated students taking math courses, AY 2009−10 to 20014−15
Number of duplicated students
40000
30000
20000
10000
0
920
950
990
1010
1020
1030
1040
1050
1060
1210
Math course


Duplicated students taking math courses by year
Number of duplicated students
8000
6000
4000
2000
0
AY_2009−10
AY_2010−11
AY_2011−12
AY_2012−13
AY_2013−14
AY_2014−15
year
factor(course)
920
950
990
1010
1020
1030
1040
1050
1060
1210

3.1 Course performance

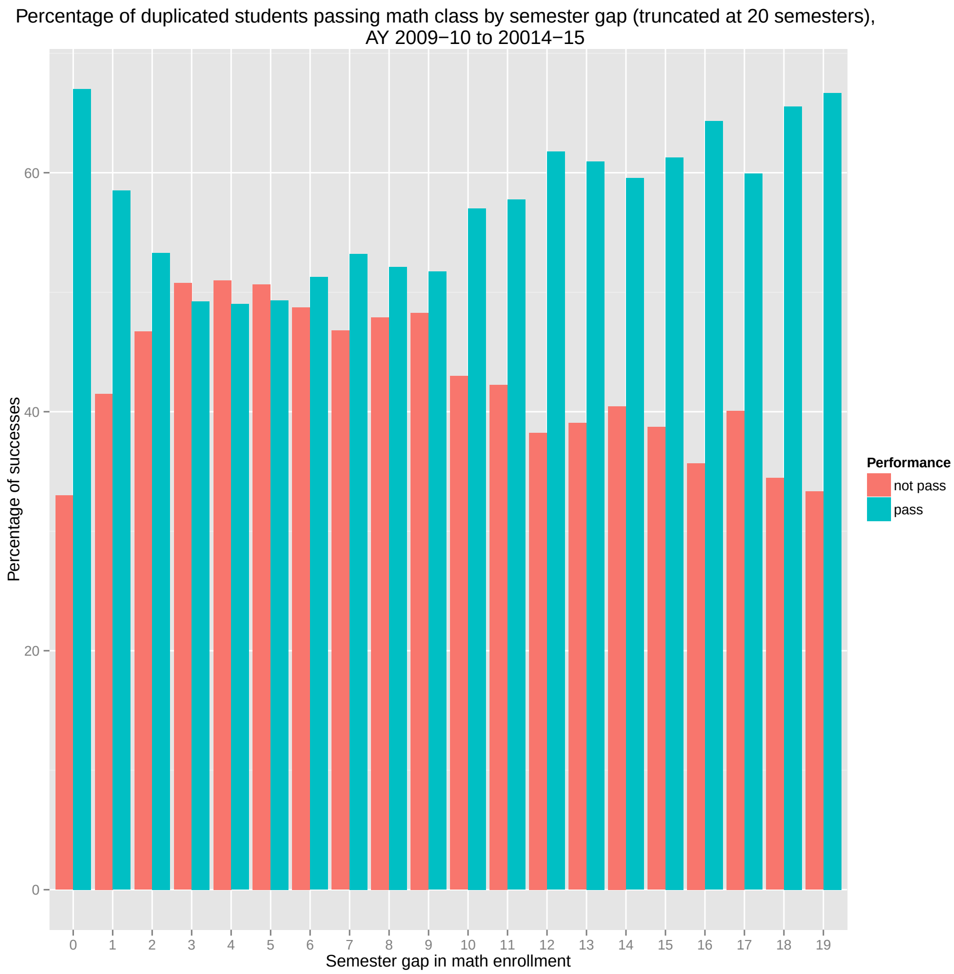




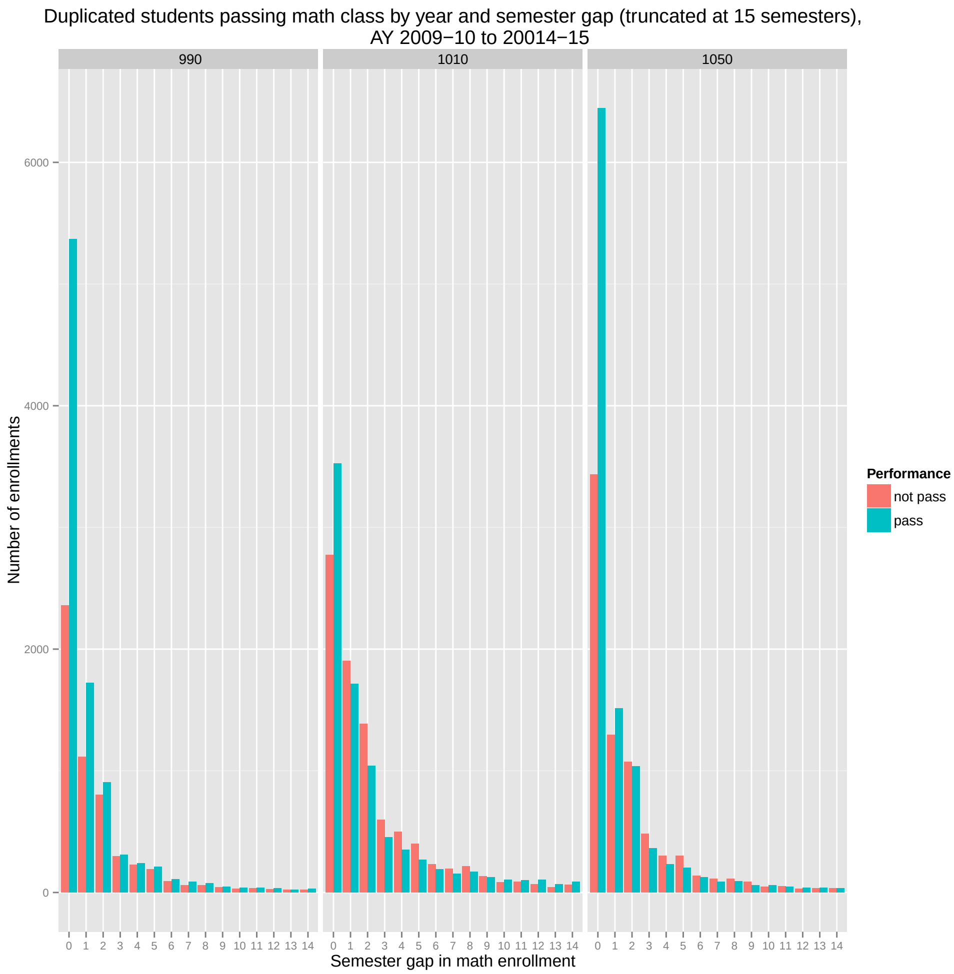
The above plot shows counts of passing/failing enrollments by semester gap since last math class. Students are represented multiple times in this graph, once for each time they take a class after their first. Note that these counts are not corrected for student ability or teacher influence. Hence the relationship between gap and passing rate should not be interpreted as a causal one.

Observations:

- The number of successes outnumbers failures for gaps of 0, 1 and 2 semesters.
- For gaps of 3, 4, and 5 the number of failures slightly exceeds the number of successes.
- For gaps of length 6 and larger the number of successes again exceeds failures—surprisingly with a margin that increases with the size of the gap.



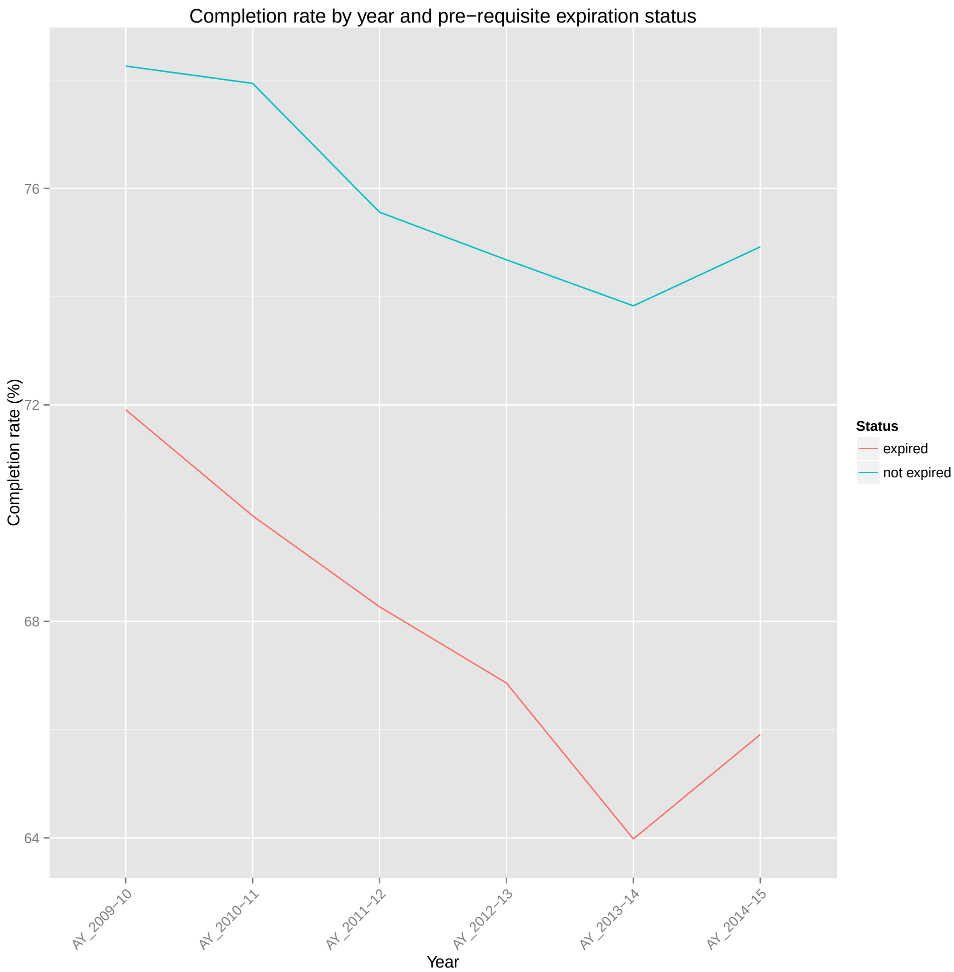
This graph shows the same relationship as in the previous graph, but with the y-axis defined by the percentage of successes within each increment in enrollment gap. (That is, the two bars for each gap increment add up to 100.) While the passing rates on the right hand side (for large gaps) resemble those on the left hand side, it should be kept in mind that the numbers involved are relatively low.

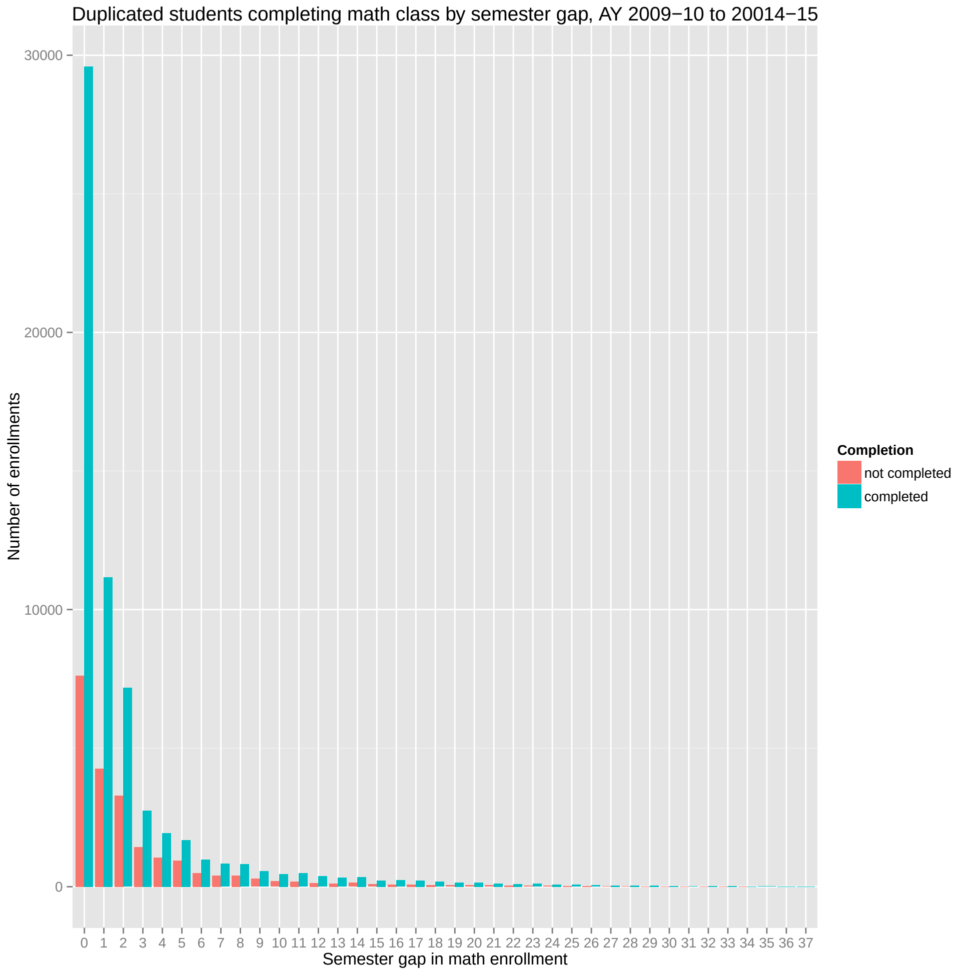


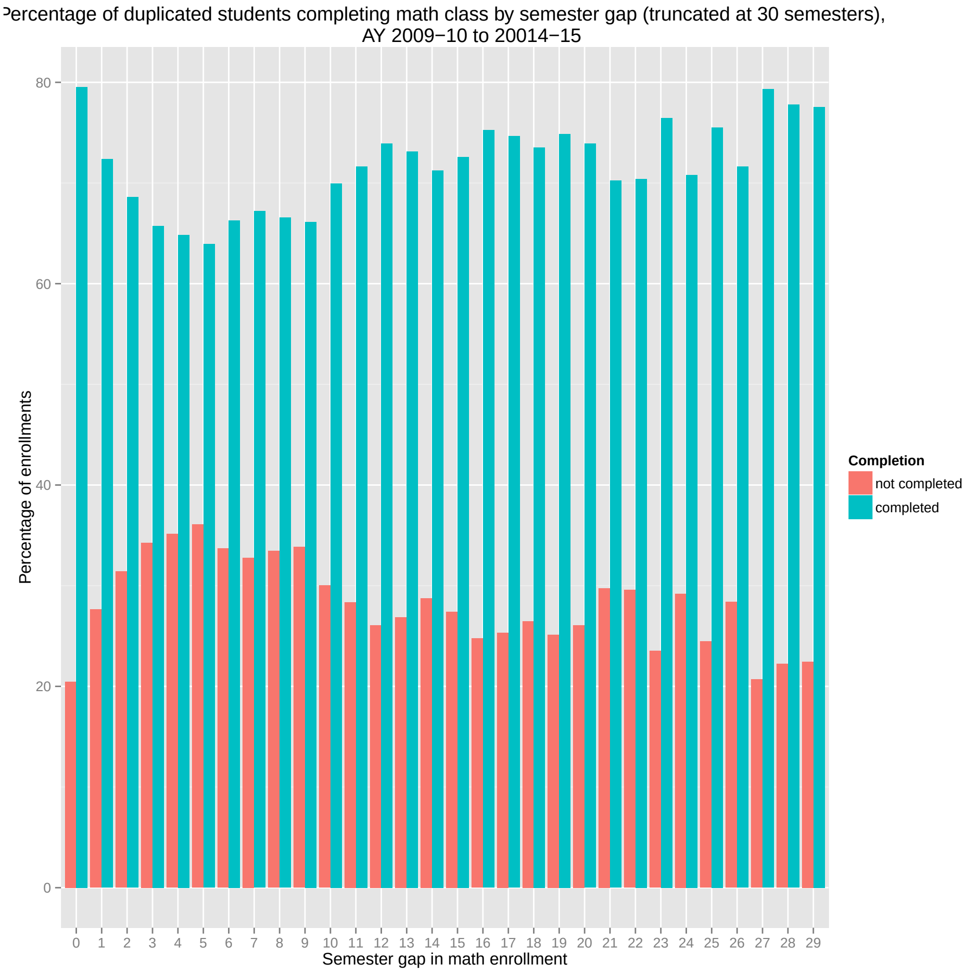
Observations: Pass rates decline more sharply as the gap increases for 1010 and 1050 than for 990. Failures outnumber successes by gap = 1 in 1010 and by gap = 2 in 1050.

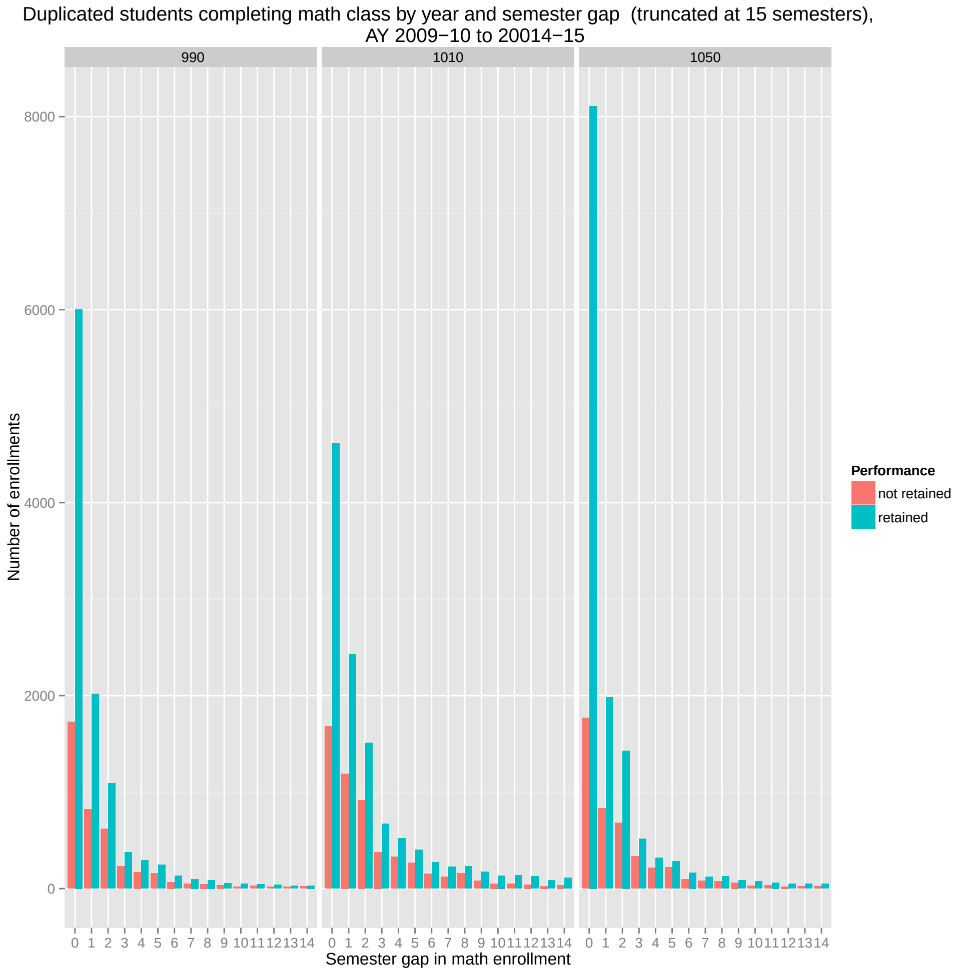
3.2 Course completion

Completion is defined by a grade of D- or above. Grades of E, W, or (rarely) I have been coded as not complete.


Completion rate by year and pre−requisite expiration status
Completion rate (%)
76
72
68
64
AY_2009−10
AY_2010−11
AY_2011−12
AY_2012−13
AY_2013−14
AY_2014−15
Year
Status
expired
not expired


Duplicated students completing math class by semester gap, AY 2009−10 to 20014−15
Number of enrollments
30000
20000
10000
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37
Semester gap in math enrollment
Completion
not completed
completed


Percentage of duplicated students completing math class by semester gap (truncated at 30 semesters),
AY 2009−10 to 20014−15
Percentage of enrollments
Semester gap in math enrollment
Completion
not completed
completed
0
20
40
60
80
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29



4 What if? Estimating effect sizes

How would enforcing the expiration on math pre-requisites impact course performance/completion? This is a question about effect size. It is worth thinking about effect size before getting into statistical modelling, since statistically significant effects sometimes aren't large enough to be practically meaningful. So, let's assume for the sake of argument that we've found statistically significant effects—found a causal relationship between pre-requisite status (expired, not expired) and rates of course performance/completion—such that enforcing expirations would have a positive impact on rates. How big might that impact be?

We can use the following conceptual model to do a rough, back-of-the-envelope estimate of the upper and lower bounds on effect sizes.

- *Upper bound.* Suppose that students with expired pre-requisites were blocked from enrolling, and as a consequence did not take a math course. Then the pass/completion rate in any given year would simply be the existing rate for students with non-expired pre-requisites. If some of students with expired pre-requisites used the accuplacer to enroll in a math class, then overall enrollment would go up over the previous scenario but pass/completion rates probably would not. That is, such students might do as well as those with non-expired pre-requisites, but there is no reason to think they would do better.
- *Lower bound.* Suppose that students with expired pre-requisites were blocked from enrolling, and as a consequence used the accuplacer to qualify for, and take, a math course. Then the pass/completion rate in any given year would be the same as, or possibly better than, the existing rate for all students,

expired and non-expired combined, but certainly no worse. In this case, enrollment would stay the same as it was before enforcement. Essentially, this is a business-as-usual scenario.

If expirations were enforced pass/completion rates should thus lie somewhere in the range defined by an upper bound equal to the existing rate for students with non-expired pre-requisites and a lower bound equal to the existing rate for all students.

4.1 Upper and lower bounds for pass rates

Here is a table summarizing the upper and lower bounds by year for pass rates using the above model:

year	lower_bound	upper_bound	range
AY_2009-10	65.21	66.58	1.37
AY_2010-11	63.98	65.61	1.63
AY_2011-12	59.46	61.55	2.09
AY_2012-13	59.76	62.32	2.56
AY_2013-14	56.33	58.93	2.60
AY_2014-15	57.16	59.95	2.79

The fourth column in this table, range, is the difference between the upper and lower bounds. It defines the maximum effect for each year. If expirations had been enforced in the 2009-10 academic year, for example, we would expect pass rates to have ranged between 65.21 and 66.58. In that case the effect size would have been somewhere between 0%, if the pass rate had come in at the lower bound, and 1.37%, if it had come in at the upper. (By "effect size" in this case we mean the percentage change in the pass rate as a result of enforcing expirations.)

What sort of effect size should we expect in future years? Very roughly—our intention here is just to give a sense of the scale of the effect—we could say that the minimum effect size would remain at 0%, as in the case above, while the maximum might be as high as the maximum we've seen historically: 2.79%.

4.2 Effect sizes for completion rates

Here is a table summarizing upper and lower bounds by year for completion rates:

year	lower_bound	upper_bound	range
AY_2009-10	77.04	78.26	1.22
AY_2010-11	76.30	77.94	1.64
AY_2011-12	73.86	75.56	1.70
AY_2012-13	72.51	74.68	2.17
AY_2013-14	71.26	73.83	2.57
AY_2014-15	72.44	74.92	2.48

Using the same logic as above: if expirations were enforced we would expect that, based on historical data, effect sizes for future completion rates would have a minimum of 0% and a maximum that could go as high as the maximum we've seen historically, 2.57%.