Student Success in Math at Los Angeles Mission College: The Effect of Placement and Early Online Class Performance

Matthew Del Castillo, Tracy Christine Albers
Los Angeles Mission College*

May 2019

Abstract

In an attempt to evaluate the success of math students at community college, we evaluated the success of students placed into higher courses and also evaluated student success in an online math course. We addressed two questions in this research: (1) How did the Multiple Measures Assessment Project affect student success? and (2) Can a student's success be predicted in the first two weeks of an online math course.

To evaluate the Multiple Measures Assessment Project, we isolated two sets of data for similar students before and after the implementation. We evaluated totals for students attempting courses, passing courses, passing higher level courses, and whether the groups as wholes were more likely or less likely to achieve success within four semesters. To predict student success in an online course, we evaluated data in the first two weeks of a semester and used various logistic regressions on different factors to evaluate the likelihood that a student would receive a passing grade at the end of the semester.

Our research indicates that the students who were received a higher placement due to the implementation of the Multiple Measures Assessment project were more likely to achieve success within four semesters compared to the group unaffected by the implementation. We were also able to predict a student's success based on early semester performance with different models with accuracies between 60 and 80 percent.

This research could indicate that a student allowed to take higher level math courses is more likely to succeed in community college than students not given the opportunity. The success rates of early online class is important for identifying students unlikely to pass the course. This information can be used by students, professors, and college faculty in evaluating the best course of action for that student and the class.

*Partially Supported by NSF-DMS-1541911 RE-C^2 grant. RE-C^2: Research Experiences in Community Colleges.

1 Math Placement

1.1 Introduction

The purpose of the first project was to study the effects of the Multiple Measures Assessment Project on student success in Math Courses at Los Angeles Mission College. Success was defined as passing a 200 series course within four semesters. We also studied success per course and specific success in Statistics (Math 227) as well.

1.2 Background

Prior to the Spring Semester of 2017, the majority of students who began attending Los Angeles College would be placed into an appropriate math course based on the results of an assessment test known as the Math Diagnostic Testing Project. However, starting in Spring 2017, students could also complete a questionnaire called the Multiple Measures Assessment Project. This questionnaire involved questions about the highest math course taken during High School, the grade received, and other questions pertaining to study habits. These responses were reported by the students and not subject to verification. Based on the answers provided by the students, a student could potential be placed up to two courses ahead of where they had been placed using the more traditional placement test.

1.3 Methodology

We received two data sets from the Los Angeles Mission Math Department related to Project 1. The first was data from students who registered and assessed starting in Spring of 2017. This data was organized by student and included the semester of their assessment, their MDTP placement data, their MMAP placement data, and their Final Placement. The data also included all math courses taken at any of the nine community colleges in the Los Angeles Community College District and the grade they received in the course. This data extended until Summer of 2018.

The second data set was organized by math course taken and had an identifier, a 4-digit code indicating the semester and year, the course taken and the grade achieved. There was a second document that contained the students assessment data organized by the semester they assessed. It should be noted that students did not take courses the semester they assessed and typically began courses the following semester.

To maximize the usable data and to achieve more desirable results, we limited the data to two major groups each consisting of two smaller data sets. Pre-MMAP 1 included students who assessed in Spring of 2016, began classes in Summer of 2016, and their progress until the end Spring of 2017. Pre-MMAP 2 included students who assessed in Summer of 2017, began classes in Fall of 2016, and their progress until the end of Summer 2017. Post-MMAP 1 included students who assessed in Spring 2017, began classes in Summer of 2017, and their progress until the end of Spring 2018. Post-MMAP 2 included students who assessed in Summer of 2017, began classes in Fall of 2017, and their progress until Summer of 2018. Both Post-MMAP data sets focused specifically on student's whose Final Placement reflected a bump from their MMAP questionnaire that exceeded their MDTP placement.

Pre-MMAP 1 and Pre-MMAP 2 totals were then combined to compile the results for Pre-MMAP and similarly Post-MMAP 1 and Post-MMAP 2 were combined to compile the results for the Post-MMAP students who took a bump in placement. This resulted in two populations who attended community college in similar years under similar conditions for the same number of semesters with the only known difference between the two populations being the choice to take a bump in placement.

Using R, totals were gathered from all of the data sets. Categories that were counted were: number of attempts at a course by any student, final attempts at a course by a student, and final number of students who passed a specific course. This was done for nineteen math courses. Note that the number of attempts at a course by any students counts every attempt at a course even the attempt was made by a student more than once, conversely, the final attempt only counts one attempt per student even if that student attempted that specific math course more than once.

Additional calculations were done for 200-series courses. Specifically, totals were gathered for all attempts at any 200-series course, total attempts per student at any 200-series course, and the total number of students who passed at least one 200-series course. We used these totals to evaluate the the students who passed a 200-series course as a percentage of their cohort in order to evaluate the overall success of the Pre-MMAP cohort and the Post-MMAP cohort.

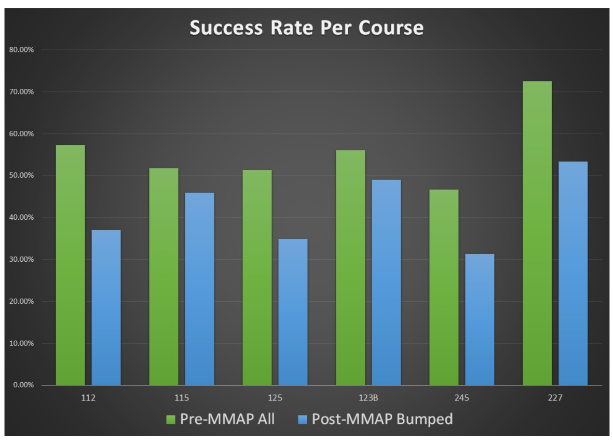
Finally, we performed calculations on Statistics (Math 227). We chose to focus on Statistics because it had a large pool of students across the different data sets, it has a dedicated path to success wherein a student takes Pre-Statistics (Math 137) in lieu of the Algebra series and then is better prepared to pass Statistics. Statistics is also one of the more accessible for students as means of achieving success in any 200-series course. We evaluated Statistics in two different ways. (1) What is the likelihood that a student will pass Statistics on the condition that the student previously passed Pre-Statistics and (2) What is the likelihood that a student will pass Statistics if they were bumped into Statistics with the MMAP assessment when their MDTP assessment was for a lower division class.

1.4 Results

This chart contains the totals for each math course evaluated. Total Attempts: Attempts at a course by all students including repeated attempts; Final Attempts: Attempt per student; Passed(A-C): Number of times a course was passed by a student. The success rate was calculated as the ratio of Passed(A-

PreMMAP Total Students: 1520					Post MMAP (Bumped) Total Students: 638				
Course	Total Attempts	Final Attempt	Passed (A-C)	Success Rate	Total Attempts	Final Attempt	Passed (A-C)	Success Rate	Course
105	42	40	20	50.00%	3	3	2	66.70%	105
112	229	205	119	58.00%	57	46	17	37.00%	112
115	459	390	170	43.60%	167	133	61	45.90%	115
125	329	287	152	53.00%	103	86	30	34.90%	125
137	33	31	19	61.30%	9	9	5	55.60%	137
123A	64	58	30	51.70%	19	17	7	41.20%	123A
123B	52	47	27	57.40%	57	49	24	49.00%	123B
123C	31	31	22	71.00%	12	12	10	83.30%	123C
121	6	4	3	75.00%	2	2	0	0.00%	121
215	2	2	1	50.00%	1	1	0	0.00%	215
245	84	68	34	50.00%	82	67	21	31.30%	245
238	5	5	2	40.00%	7	6	3	50.00%	238
227	112	100	69	69.00%	244	208	111	53.40%	227
240	28	26	16	61.50%	46	41	31	75.60%	240
260	9	9	6	66.70%	11	10	7	70.00%	260
265	16	15	10	66.70%	14	13	10	76.90%	265
266	4	3	1	33.30%	7	6	3	50.00%	266
267	1	1	1	100.00%	1	1	0	0.00%	267
200*	261	195	140	71.80%	425	313	190	60.70%	200*

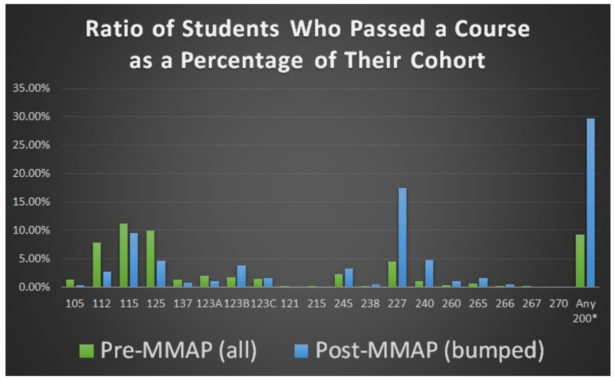
C) / Final Attempt. Note that *200 indicates an attempt at a single 200 series course.

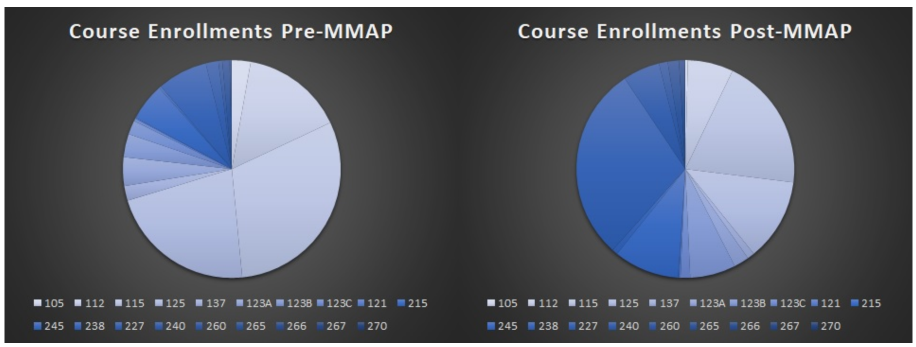


In this graph, we display the Success Rates Per Course. Specifically, these are courses with more than 50 students in both the Pre-MMAP and Post-MMAP groups. This chart reveals an apparent decline in Success Per Course after the implementation of MMAP. However, in the graph below, we have calculated the ratio of students who pass a course as a percentage of their cohort. Here we see that a higher percentage of the Post-MMAP group succeeded in passing course 123B, 123C and almost all 200-series courses.

These two pie charts indicate the number of course enrollments in the Pre-MMAP and Post-MMAP cohorts. The pie chart on the left shows that majority of the Pre-MMAP students enrolled in lower division courses. The pie chart on the right shows that a much higher percentage of students in the Post-MMAP cohort enrolled in higher division courses over a similar period of time.

While evaluating for success, we paid special attention to Statistics (Math 227). Using the Pre-MMAP dataset, we calculated the likelihood that a student



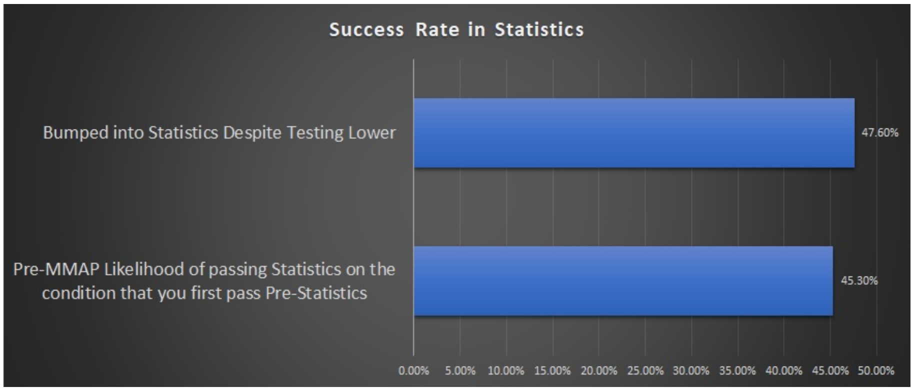


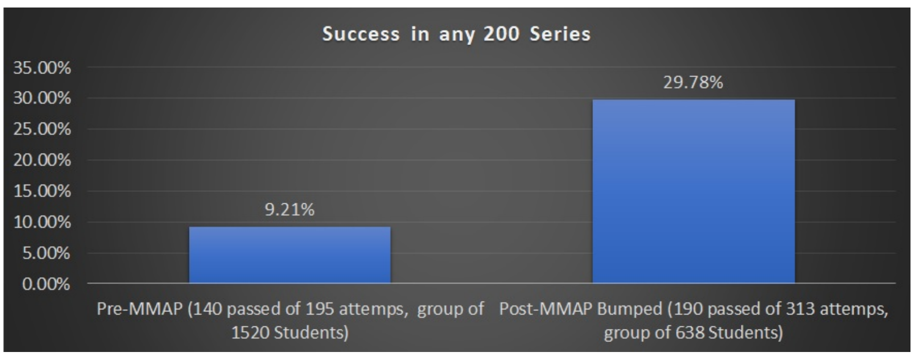
would pass Statistics on the condition that they first completed Pre-Statistics (Math 137). We then took the product of this likelihood and the independent likelihood of a student passing Pre-Statistics resulting in a calculated likelihood of passing of 45.30 percent. We then investigated the Post-MMAP data set for students who were bumped into Statistics despite testing lower initially. These students had a 47.6 percent likelihood of passing.

We defined success as the completion of a single 200-series course within 4 semesters. Therefore, we evaluated the percentage of successful students in each cohort. The Pre-MMAP cohort consisted of 1520 students. Of those students, 195 attempted any 200-series course and 140 student passed a 200-series course. The percetage of students who passed a 200-series course of the Pre-MMAP cohort is 9.21 percent. The Post-MMAP cohort studies consisted of 638 students. Of those students, 313 attempted a 200-series course and 190 passed a 200-series course. The percentage of students who passed a 200-series course of the Post-MMAP cohort is 29.78 percent.

1.5 Limitations

After the relevant data was sorted and categorized, it was noted the the final data groups were fairly small and this limits the reliability of the results. This is especially true for success rates in courses with only a handful of students.





The Post-MMAP cohort consisted of students whose final placement was greater than their MDTP placement indicating that they reported higher skills and then chose to take a more advanced class. These students were likely more confident in their mathematical and academic skills. This could be reflected in their overall performance during the time studied.

1.6 Conclusion

Our results indicate that the success rates per course dropped notably after the implementation of MMAP. However, there appeared to be a boost in throughput with a higher percentage of students achieving success in the four semesters.

2 Early Online Class Performance

2.1 Introduction

The purpose of this second project was to predict individual student success in Math 125 (Intermediate Algebra) online classes at Los Angeles Mission College during the Fall 2018 semester, using their performance in the first two weeks. Success was simply classified as a grade of C or higher in the course at its conclusion.

2.2 Background

Math 125 is a lower division intermediate algebra course offered within the Los Angeles Community College District. Interactions between students and instructors are limited to online discussions through the Canvas Learning Management System, emails, physical office hours on campus, and in-person exams. The bulk of the course content and assignments are administered through MyMathLab, a Pearson product that provides a platform for homework assignments, quizzes, and eTexts to be given to students entirely online. Students are given a 14 day grace period before committing to spending money to access the course content, after which they must pay 95 dollars to continue.

The platform tracks students' grades automatically, along with a few other factors that are available to instructors, like time spent on individual assignments, and timelines of completed assignments.

2.3 Methodology

We received a combined and re-coded data set from multiple instructors at Los Angeles Mission College for a total of 333 students in Fall 2018 Math 125 courses. It included each student's recoded ID, overall score, median time spent on assignments, total time spent on assignments, and the score, time spent, and last date worked for all assignments assigned over the duration of the course. This amounted to 192 total features in the data set.

In Excel, the data was pruned of overall scores and replaced with a pass/fail condition for each student and missing assignments and fields were filled in with zeroes. All further work with the data was done in R after this point. Dates were originally in hh:mm:ss format, so they were converted into decimal day values for easier comparison and processing through machine learning algorithms. Last dates worked for each assignment were trimmed from the data set to focus on the individual assignment scores and time spent on each.

In order to measure effectiveness of various combinations of factors in a more controlled manner, we decided on using only one algorithm for attempting classification. Logistic regression was the ideal choice, because it provided an easy platform for classification, and generated probabilities for pass or fail to generate more nuanced results in testing multiple probability thresholds.

Combinations of factors that were of interest were gathered, including the primary focus of the first two weeks of assignments. Single factors were also tested to provide a baseline for effectiveness of only one variable. Factors such as total time spent in the course, median time spent, and the students' score on a final review assignment were considered for their own models.

Each model was trained on a random 75 percent split of the original 333 observations, with the other 25 percent reserved for testing afterwards. In addition to intuitive combinations, both forward and backward stepwise iterators were applied to base models with one or all factors respectively, to determine if there were important factors that we were overlooking in our initial assessment of the data set. Their performance was rated on the final iteration of each.

To gauge performance, we used 2 groups of metrics. The first is the standard confusion matrix for classification, with threshold for prediction set at 0.5 for all models. The confusion matrix creates a two-way table for actual classification vs predicted classification. It includes 4 different numbers:

$$\begin{bmatrix} TN & FN \\ FP & TP \end{bmatrix}$$

TN for true negative, meaning a student who failed was correctly classified as failing. FN for false negative, meaning a student who passed was incorrectly classified as failing. FP for false positive, meaning a student who failed was incorrectly classified as passing. TP for true positive, meaning a student who passed was correctly classified as passing.

Using these 4 measures, we calculated 5 metrics:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

ACC, or accuracy, measures the proportion of correct predictions out of all observations.

$$TNR = \frac{TN}{TN + FP}$$

TNR, or true negative rate, measures the proportion of correctly predicted failing students out of failing students..

$$FNR = \frac{FN}{TP + FN}$$

FNR, or false negative rate, measures the proportion of incorrectly predicted failing students out of passing students.

$$TPR = \frac{TP}{TP + FN}$$

TPR, or true positive rate, measures the proportion of correctly predicted passing students out of passing students.

$$FPR = \frac{FN}{TN + FP}$$

FPR, or false positive rate, measures the proportion of incorrectly predicted passing students out of failing students.

The second metric used for gauging performance was the ROC (Receiver Operator Characteristic) curve, which plots the false positive rate vs the true positive rate of prediction at every threshold, instead of just the static 0.5 threshold used for the confusion matrix. This provides us with a different perspective of overall model performance. An ideal ROC curve for a model that predicts things correctly at every threshold would go along the left side vertically, and then go to the right.

The sister metric to the ROC curve is the AUC (Area Under the Curve), which measures the area under the area under the ROC curve. Perfect model predictions would lead to an AUC of 1, and completely incorrect model predictions would lead to an AUC of 0. Both of these metrics were calculated and plotted for all models.

2.4 Results

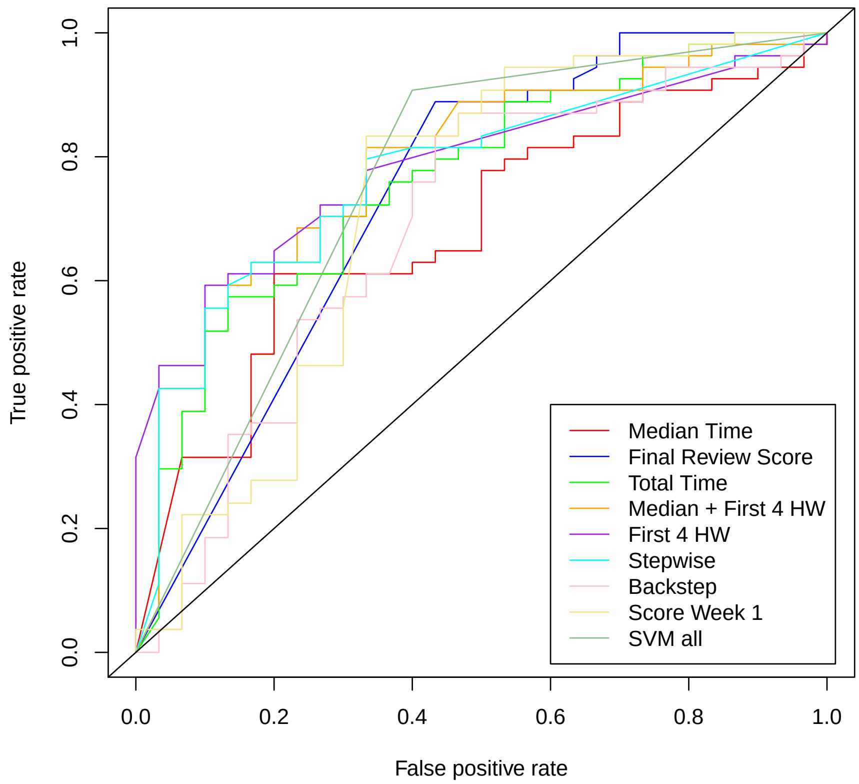
Numerical metrics for all 9 models were calculated and conglomerated for comparison.

Predictor	ACC	TNR	FNR	TPR	FPR	AUC
Median Time Spent	0.63	0.93	0.07	0.10	0.90	0.68
Final Review Score	0.74	0.89	0.11	0.47	0.53	0.73
Total Time Spent	0.74	0.89	0.11	0.47	0.53	0.76
Median + First 4 HW	0.71	0.70	0.30	0.73	0.27	0.79
First 4 HW	0.71	0.72	0.28	0.70	0.30	0.79
Stepwise	0.71	0.70	0.30	0.73	0.27	0.77
Score 1,4,9+Time 1,6,8,40	0.74	0.83	0.17	0.57	0.43	0.68
Pass By Week 1	0.74	0.85	0.15	0.53	0.47	0.72
SVM All Factors	0.80	0.91	0.09	0.60	0.40	0.75

After comparing 2 families of metrics, all models had accuracies ranging between 60 and 80 percent, with AUCs of .7 to .8. By accuracy, single factor logistic regressions ran on final review score, total time spent on assignments, and week 1 passing all scored .74 in accuracy, along with the backwards stepwise regression. They all had similar scores for the other 4 confusion matrix metrics as well. Positive classes seem to be overrepresented in predictions, as FPR for all 4 is near or above 50 percent.

Looking at AUC values, the two models including the first four assignments did the best, with AUCs of .79. They had more balanced metrics for confusion matrix prediction rates. Overall accuracy was only marginally lower than the 4 highest scoring models, at .71 compared to the aforementioned .74.

Comparing the ROC curves for all models, there are no clear best predictors, as they seem to group within the same band. Median time spent seems to show



up as a marginally worse predictor based on its ROC curve that is below the rest, and is reflected in its AUC value of .68. As this proves visually inconclusive, we fell back on interpreting AUC values for each model as they provided a clearer division between model performances.

Overall, although the models trained with the first two weeks of assignments seemed to do the best considering both confusion matrix metrics and AUC, the difference among all 9 is fairly marginal, and could potentially be attributed to randomness in the test set.

2.5 Limitations

Some problems present themselves readily at the conclusion of this research. For one, an additional data set for Summer 2018 Math 125 courses was also provided, but was outside the scope of the study. It was also an even smaller data set, only amounting to about 80 observations. However, it included another facet of online participation: Canvas discussions and miscellaneous assignments that were not listed through MyMathLab. In further research this could prove useful for training a more accurate model.

Since the classification was based on overall score, a large portion of the features in the data set ended up being dependent, as calculating the average of all assignment scores would give you the students' passing or failing grade 100 percent of the time. Lack of diversity in features was also an issue, and did not provide much variety for testing different model combinations for outstanding features. Generated features were limited to score thresholds after certain points in the class.

2.6 Conclusion

Our results indicate that while the first 2 weeks of assignment scores and times spent make for decent predictors of passing an online Math 125 course with an AUC value of .79 under logistic regression, they only do marginally better than other combinations of factors such as median time spent on assignments, and all assignments combined when comparing prediction rates.

Acknowledgement

We would like to thank our research advisor and professor Emil Sargsyan, who helped us through this entire project and arranged much of the data collection. We would also like to thank the LAMC math department for their cooperation in providing data and other information related to the project. In addition, our advisors at CSUN, Bruce Shapiro and Adriano Zambom were instrumental in getting us on our feet in the machine learning and data science realm. Finally, we would like to thank Helena Noronha for allowing us the opportunity to participate in the RE-C^2 program and conduct this research in the first place.