Multiple Measures Assessment Project (MMAP)
Summary of Methodology for Phase II Rule Sets and AB 705 Adaptations
English and Math

Introduction

The goal of the original MMAP placement recommendations was to identify students who were highly likely to succeed at any given level of an English and math course. The results suggested that many students had been placed too low, with underrepresented minority and female students placed disproportionately lower than their non-minority and male peers, especially in math.

The passage of AB 705 required the field to think differently and determine how best to maximize overall completion of the transfer-level course ("throughput") in one year rather than determining who is highly likely to succeed in a particular course, regardless of level. The throughput rate is the percentage of students who complete transfer-level English or math courses with a grade of C or better within two primary semesters or three primary quarters. For students beginning below transfer-level, the throughput rate is defined as the percentage of students who attempt and succeed in the remedial course in which they are placed, then enroll and pass the subsequent remedial courses, before finally enrolling in and successfully completing the first transfer-level course. Rather than identifying students who are highly likely to succeed at a particular level, AB 705 changes the task to identifying students highly unlikely to succeed if directly placed into transfer-level.

MMAP Methodology

The MMAP decision trees include all students enrolled in an English or math course in the California Community College (CCC) system who also had four full years of high school data available in the California Partnership for Achieving Student Success (Cal-PASS Plus) data system, resulting in one million cases. As such, MMAP research is based on retrospective data—students who have already been placed, enrolled, and completed courses in the CCC system between 2008 and 2014. The decision trees depict students' success rates using course grades in the first course a student completed at a CCC and other transcript information through 11th grade for students attending community college directly out of high school and for whom 12th grade information would not yet be available for use in placement (rule sets for students who completed 12th grade are also available). Example transcript information included in the models includes cumulative unweighted high school grade point average (HSGPA), highest course completed, type of course completed (by level and subject matter), and delay between high school and college.

Decision trees were selected due to their popularity as a machine-learning tool and the ease of interpreting and understanding results. Decision trees are a form of data modeling that results in a set of "if-then" rules, where the if-then statements are referred to as "nodes" or "leaves" of the tree. Decision rules were formed by setting certain criterion, so that if students met or exceeded a minimum average probability of success at a certain level, they would be grouped together. The criterion (70% predicted success rate at transfer-level) represents the minimum average successful completion rate of groups of students that the model places into that level of courses. Since the rules were based on the average minimum successful completion rate at the criterion or higher, the actual average for students placed at each level is often above the minimum threshold.

AB 705 Adjustments to the Rule Sets

In the decision trees, a machine-learning algorithm had split the data to create groups or "nodes" of students with similar success rates based on the inputs (high school transcript information). For example, students with an HSGPA below 1.9 have a 43% predicted success rate in transfer-level English, the lowest node, while students with an HSGPA greater than or equal to 3.1 have an 87% predicted success rate (the highest node). To meet AB 705 compliance, the MMAP research team focused on the lowest nodes with the lowest success likelihoods, as these students are expected to be least likely to succeed in the target course and presumably most likely to benefit from placement into developmental courses. While these lowest node characteristics involve only HSGPA, higher nodes for math courses involve other characteristics, such as completion of Algebra II.

The success rates of students in the lowest node are based upon the rates of success among students with the lowest high school GPA who were also eligible to enroll into the transfer-level course. There are several ways in which a low-GPA student could have received a transfer-level placement: scoring sufficiently high on a placement test; placement from another college with different cut scores; alternative assessments such as scores on the EAP, SAT, ACT, or AP exams; transferring coursework from other institutions; challenging the prerequisite; or an error in the intake process.

There is a concern raised from the field that direct transfer-level placement of students who both had lower high school performance and were historically placed into remediation may lower the overall success rates in the transfer-level course. To estimate the impact of this concern, referred to as "double placement," the MMAP research team focused on transfer-level English, Statistics, and Pre-Calculus, and the characteristics of the students in the lowest node. The MMAP research team obtained ACCUPLACER scores for a subset of English and math students and used multiple approaches to estimate direct placement success rates for the students in the lowest node. The multiple approaches included fitting a regression model that predicts success in target courses based on high school GPA and ACCUPLACER scores, calculating the mean high school GPA and test scores of the first attempted course, and using standard error of prediction from the regression model at each level to create lower- and upper-bound estimates (referred to as a regression adjusted model), similar to margins of error.

Conclusion

Based on the analysis described above, for all three courses, students who started directly at transfer-level completed the transfer-level course at a higher rate than students who started below transfer-level. Direct transfer-level completion in English among students in the lowest node (HSGPA < 1.9) was 43% (39% for the regression adjusted model) compared to 12% among students who started at one level below transfer-level. For Statistics, completion rates for the lowest node (HSGPA < 2.3) were 40% (29% for the regression adjusted model), compared to 10% for students who started at one level below; for Pre-Calculus, completion rates in the lowest node (HS GPA < 2.6) was 38% (27% for the regression adjusted model) compared to 13% for students who started at one level below. This analysis was unable to provide evidence that students with a low HSGPA would have higher throughput rates by being placed into traditional remediation; therefore, we were unable to identify any group of students highly unlikely to succeed if directly placed into transfer-level English, Statistics, or Pre-Calculus. It is recommended that local researchers compare below-transfer-level remediation throughput rates to direct transfer-level placement success rates both with and without specialized support such as co-requisites. Additional detail on the Phase II rule sets can be accessed here.