# Improving Placement Accuracy in California's Community Colleges Using Multiple Measures of High School Achievement

Peter Riley Bahr*
University of Michigan

Loris P. Fagioli
Research and Planning Group for California Community Colleges

John Hetts
Educational Results Partnership

Craig Hayward
Bakersfield College &
Research and Planning Group for California Community Colleges

Terrence Willett
Research and Planning Group for California Community Colleges

Daniel Lamoree
Educational Results Partnership

Mallory A. Newell
De Anza College &
Research and Planning Group for California Community Colleges

Ken Sorey
Educational Results Partnership

Rachel B. Baker
University of California, Irvine

*Correspondence may be directed to the first author at the Center for the Study of Higher and Postsecondary Education, University of Michigan, 610 E. University Avenue, Ann Arbor, MI 48109-1259; email: prbahr@umich.edu. This paper was presented April 29, 2017, at the annual meeting of the American Educational Research Association in San Antonio, Texas. The research discussed here is a result of a collaboration of the California Community Colleges Chancellor's Office, California Community Colleges Common Assessment Initiative, the Research and Planning Group for California Community Colleges, and Educational Results Partnership, and it received funding from the Butte-Glenn Community College District. The authors gratefully acknowledge contributions and feedback on this work from Alyssa Nguyen and Nathan Pellegrin.

**Improving Placement Accuracy in California's Community Colleges Using Multiple Measures of High School Achievement**


## ABSTRACT

Standardized placement tests remain the primary means by which new community college students are assessed and placed in the hierarchy of math and English coursework, from developmental to advanced college- or transfer-level courses. However, a growing body of evidence indicates that such tests tend to underestimate students' likelihood of succeeding in college-level coursework, leading to students being misplaced in developmental coursework, slowing their academic progress and increasing their likelihood of dropping out of college. Here, we discuss the results of a statewide research effort to improve the accuracy of student placement in math and English in the California Community College system. We apply decision tree methods to high school and college transcript data to identify key measures of high school achievement that predict success in nine levels of community college math courses (three developmental and six college-level) and four levels of community college English courses (three developmental and one college-level). Drawing on our findings, we develop and validate placement rule sets that colleges can apply directly in their placement processes. We then discuss challenges in communicating the results to community college stakeholders, implications of the results for institutional planning and practice, and important directions for future research on this subject.


*Keywords:* community college, high school, achievement, assessment, placement, multiple measures, developmental, remedial, decision tree, data mining

**Improving Placement Accuracy in California's Community Colleges Using Multiple Measures of High School Achievement**

## INTRODUCTION

Nationwide, a majority of first-time students in community colleges are placed in developmental coursework in math, English, or both, and few of these students successfully navigate the curriculum to complete college-level courses in these subjects (Bailey, 2009). However, evidence indicates that a meaningful percentage of students are assessed as needing developmental instruction despite being likely to succeed in college-level coursework, resulting in them being under-placed in the curriculum (Belfield & Crosta, 2012; Scott-Clayton, 2012a). Consequently, researchers, policymakers, and college administrators increasingly are focusing their attention on improving assessment and placement processes to more accurately determine students' capabilities and avoid unnecessary skill remediation with its corresponding risk of premature departure from college (Scott-Clayton, Crosta, & Belfield, 2014a).

A number of studies have examined the use of alternative or supplementary information to more accurately place community college students in math and English. These studies generally indicate that a student's high school achievement provides predictions of course outcomes in math and English that are superior to predictions based solely on placement exam scores (Bahr, 2016; Ngo & Kwon, 2014; Scott-Clayton et al., 2014a). Thus far, however, research on this subject has shed little light on the particular measures of high school achievement that are the best predictors of course outcomes, how these predictors differ between math and English, and how they vary across the levels of skill represented in the coursework that composes the curriculum of each subject (e.g., statistics versus college algebra versus pre-calculus). Moreover, prior research provides only limited information about how these measures should be applied when placing students in math and English courses, particularly the thresholds on measures above which success in a particular course would be expected.

In this study, we use a data mining technique known as decision tree analysis to investigate these issues, drawing on statewide high school and community college data from California. From a range of high school achievement measures, cumulative high school grade point average (GPA) emerges as the most consistently useful measure for identifying students who are likely to succeed in a given math or English course. Drawing on the results of our analyses, we develop placement rule sets that reduce the risk of student under-placement relative to standardized placement exams and that are easily applied by community colleges in a variety of assessment and placement systems. We validate these rule sets with two separate testing samples.

## BACKGROUND

### Access versus Success

Increasing access to higher education has been a foundational principle of community colleges from their beginning over 100 years ago. To that end, most community colleges exercise relatively non-selective admissions, welcoming nearly all students who apply (Bahr & Gross, 2016). This open-door policy has created a contradiction for these institutions, however, because the majority of students who enroll are deemed to be underprepared for college-level coursework or have other circumstances that increase their likelihood of leaving college without achieving their educational goals (K. L. Hughes & Scott-Clayton, 2011; Juszkiewicz, 2014; Parker, Bustillos, & Behringer, 2010). It perhaps is not surprising then that the rates at which community college students complete postsecondary credentials or transfer to four-year institutions tends to be low, as compared with the graduation rates of four-year institutions.

In recent years, scholars have questioned whether a promise of access to higher education, when combined with a comparatively small probability of completing college, can be considered access at all (Bahr & Gross, 2016; Shulock & Moore, 2007). This shift in discourse concerning community colleges has resulted in a sizeable effort by scholars, administrators, policymakers, foundations, and other stakeholders to identify the institutional policies and practices that contribute to the loss of so many students along the path to a college degree. Because of this effort, it now is clear that many students become mired in developmental curricula in math, reading, or writing and never reach college-level coursework in these subjects, foreclosing educational opportunities in many fields of study.

**Developmental Education**

Developmental education, sometimes referred to as remedial or basic skills education, has been the most common response of community colleges to the perceived shortfalls in academic skills among newly admitted students (Scott-Clayton et al., 2014a). Nationally, over two-thirds of community college students enroll in at least one developmental course (Bahr & Gross, 2016), but an even larger percentage of community college students are placed into developmental coursework, with some students electing not to enroll (Bailey, Jeong, & Cho, 2010). Thus, the considerable majority of community college students are affected by developmental education, either through the time and cost of taking developmental courses or by being prevented from enrolling (or advised not to enroll) in college-level coursework because they have not completed the prerequisite developmental courses.

Developmental education frequently is organized as sequences of increasingly advanced courses in a subject. For example, the lowest developmental math course often is arithmetic, followed by pre-algebra, beginning algebra, and intermediate algebra (Bahr, 2012). The lowest college-level courses in math that are offered by community colleges usually include statistics, college algebra, and one or more general math courses, such as math for liberal arts majors (Bahr, Jackson, McNaughtan, Oster, & Gross, 2017). Students who are assessed as needing developmental education in a subject are placed at a particular level in the sequence and progress through the coursework to complete the highest-level developmental course, upon which they are eligible to enroll in a college-level course in that subject.

Assessed needs for developmental education vary greatly across community college students. Some students are placed just one level below college readiness in math, reading, or writing, and need to complete only one developmental course before advancing into college-level coursework in that subject. Other students are placed two or more levels below college readiness in one, two, or all three subjects and may require a year or more to complete the developmental sequences.

Developmental courses usually are presented by colleges as credit bearing (i.e., students earn credits in the courses), but these credits typically do not count toward a postsecondary credential. Thus, a student may spend considerable time and money in developmental education without making progress toward a degree.

Moreover, though developmental education is intended to prepare students with the reading, writing, and mathematics skills deemed necessary to succeed in college coursework, evidence regarding the effectiveness of remediation is mixed at best (Attewell, Lavin, Domina, & Levey, 2006; Bahr, 2008; Bettinger & Long, 2005). Long developmental sequences, in particular, present a substantial obstacle to students' progress and attainment, hindering rather than helping some students (Bahr, 2012). Thus, it is clear now that developmental education — a primary means through which the promise of access to higher education is intended to be fulfilled — is not always helping students in proportion to the direct and opportunity costs that it presents (Burdman, 2012). Further research is needed on the ways in which developmental education deters, rather than encourages, access and success.

**In Search of a Solution**

Efforts to improve developmental education and thereby increase community college students' chances of success have taken several forms. Among these, stakeholders have sought to redesign developmental curricula to accelerate the rate at which students who are directed to remedial coursework are able to advance into college-level coursework (Bailey et al., 2016; Burdman, 2012; Hayward & Willett, 2014; Kalamkarian, Raufman, & Edgecombe, 2015; Pearson, 2014).

Nodine and his colleagues (Nodine, Dadgar, Venezia, & Bracco, 2013) describe three models of acceleration. The first is co-requisite remediation, sometimes referred to as mainstreaming with supplemental support (Bailey et al., 2016), in which students who are assessed as needing remediation are placed directly into college-level coursework but with supplementary support courses. Evidence indicates that co-requisite remediation can improve the rate at which students complete college-level coursework in a subject (e.g. Cho, Kopko, Jenkins, & Jaggars, 2012; Jenkins, Speroni, Belfield, Jaggars, & Edgecombe, 2010). The second is developmental compression in which remedial curricula are modified to reduce the number of courses required to reach college-level coursework (Edgecombe, Jaggars, Xu, & Barragan, 2014). In math, this often is accomplished by differentiating a curricular pathway to degrees in the sciences and engineering, requiring advanced mathematics, from a pathway leading to other programs of study. The last is developmental modularization in which traditional remedial courses are replaced with “discrete learning units” (Pearson, 2014, p. 5) intended to strengthen particular competencies (Bickerstaff, Fay, & Trimble, 2016).

While it is unquestionably important to improve developmental curricula, it also is incumbent on colleges to ensure that assessment and placement processes are as accurate as possible. Of particular importance is minimizing the number of students who are deemed unprepared for college-level courses despite having the ability to succeed in these courses. Inaccurate placement of this sort is an obstacle in students' paths to credentials, elevates their risk of premature departure from college, and contributes to the sometimes unfulfilled promise of opportunity offered by community colleges (Burdman, 2012). In other words, even minimal remediation, when unnecessary, is problematic. Hence, the decision to place a student in developmental coursework should not be taken lightly and should be based on a thorough assessment of students' preparation and readiness for college-level coursework.

In that regard, community colleges in many states rely heavily on standardized placement exams, such as ACCUPLACER®, to assess students' content knowledge in math and English (Burdman, 2012; Fields & Parsad, 2012; K. L. Hughes & Scott-Clayton, 2011). Placement exams such as these have noteworthy strengths and weaknesses. On one hand, they can be largely automated, making them relatively inexpensive to administer and requiring few support personnel (Jaggars, Hodara, & Stacey, 2013; Scott-Clayton et al., 2014a; Sedlacek, 2004). Placement exams also have the appearance of consistency and even-handedness, which reinforces the perceived legitimacy of the placement process (Sedlacek, 2004).

On the other hand, placement exams provide only a narrow view of the complex concept of college readiness (Burdman, 2012). Consequently, evidence indicates that their capacity to predict performance in coursework is modest at best (Bahr, 2016). Furthermore, research shows that placement exams are more likely to underestimate students' capabilities than to overestimate them (Scott-Clayton, 2012). As a result, some students who have the capacity to succeed in college-level courses instead are placed in developmental coursework, increasing their risk of departure from college.

The systematic under-placement of students may be driven in part by the fact that students frequently are unaware of the consequences of their placement exam scores until after they have taken the exams and, thus, do not prepare adequately (Fay, Bickerstaff, & Hodara, 2013; Hodara, Jaggars, & Karp, 2012; Nodine et al., 2013). Unfortunately, evidence indicates that simply notifying students of consequential nature of the placement exams is not sufficient to improve performance (Moss, Bahr, Arsenault, & Oster, 2016).

Recognizing these weaknesses, community colleges across the country are exploring the use of alternative and supplemental measures of students' readiness for college-level coursework (Bailey et al., 2016; Burdman, 2012; Education Commission of the States, 2015). Among these measures, high school GPA has been found to be a strong predictor of success in college-level courses in math and English (Bahr, 2016; Scott-Clayton, Crosta, & Belfield, 2014b), a finding confirmed in four-year institutions as well (Geiser & Santelices, 2007; Hiss & Franks, 2014). In addition, Ngo & Kwon (2015) found math achievement in high school (e.g., highest math course completed and grade achieved) to be a useful supplement to standardized exams for placing students in the math curriculum. Hughes & Petscher (2016) mention a number of other measures of high school achievement that might be examined in future research, such as credits completed in math or English, and when particular math or English courses were taken, among others.

Measures of high school achievement are not without potential limitations, of course. They may have limited applicability for older students who have been out of high school for a number of years, students who did not graduate from high school, students who were home-schooled, and students who did not attend a U.S. high school. Nevertheless, the evidence indicates that high school academic achievement can be used to accurately place a large fraction of first-time community college students in the math and English curricula, and results in more accurate placement than do scores on standardized placement exams.

**This Study**

The body of research regarding the use of high school achievement measures to place students in the community college curriculum is growing. However, several critical questions remain to be addressed. First, it is not clear which measures of high school achievement are the most useful for predicting success in college-level coursework in math and English. We do not know, for example, whether broad indicators of academic achievement like overall high school GPA or 12th grade GPA are stronger or weaker predictors of success in college-level math than are narrower measures of academic performance, such as highest successfully completed math course in high school or grade achieved in the most recent high school math course. Moreover, we do not know if there are differences across subjects, such that the indicators that are most useful for predicting success in college-level math may differ from those that are most useful for predicting success in college-level English.

Second, the dichotomy of college-level versus developmental coursework belies the range of skill-levels represented in the curriculum that compose these categories. Developmental math, for example, can be composed of four or more sequential courses, from arithmetic through intermediate algebra. College-level math include college algebra, statistics, general education math, trigonometry, pre-calculus, calculus, etc., and these courses differ meaningfully in the mathematical competencies required for success. Currently, research on the use of high school achievement measures to place community college students cannot speak conclusively regarding how consistent are the predictors of success across the range of courses within a given subject. To illustrate with college-level math, the most useful predictors of success in statistics may differ from the most useful predictors of success in college algebra, which, in turn, may differ from the most useful predictors of success in pre-calculus, and so on. Prior research has offered only limited insights into distinctions between college-level math courses (Bahr, 2016; Scott-Clayton, 2012).

Finally, once the most useful predictors of success across the math and English curricula are identified, we still must determine the thresholds of achievement on these predictors above which success in a given course is likely. In other words, we must determine how to apply these predictors across the range of courses in a given subject. For example, if high school GPA is an important predictor of success in transfer-level English, one then must determine how high must be a student's GPA before success in this course is likely.

To summarize, continuing to advance research, policy, and practice concerning the use of high school achievement measures to place community college students in the math and English curricula requires answers to at least three major questions, which are as follow:

1. What are the useful predictors of success in developmental and college-level math and English coursework?
2. To what extent do these predictors vary across the levels of skill represented in the curriculum in each subject?
3. What are the appropriate thresholds of achievement to apply when placing students in the math and English curricula?

In this study, we seek to address these questions using a data mining technique known as decision tree analysis and drawing on high school and community college data from California.

**California Policy Context**

The California Community College (CCC) system is composed of 113 community colleges and serves approximately 2.1 million students per year, accounting for about one-fifth of all community college students in the United States (California Community College Chancellor's Office, 2017). Approximately 80% of newly admitted students in the CCC system are referred to developmental education (Mejia, Rodriguez, & Johnson, 2016), with Latino/a, African American, and low-income students being overrepresented. Evidence indicates that the majority of CCC students who enroll in developmental math or writing ultimately do not complete a college-level course in the subject (Bahr, 2012).

The decision in Romero-Frias et al. v. Mertes et al. (1988) prohibited California's community colleges from relying solely on standardized exams for student assessment and placement. Instead, the colleges must consider additional information about student academic readiness, referred to as "multiple measures" in the state guidelines (California Community College Chancellor's Office, 2011, p. 2.4). Interpretations of this requirement and the resulting placement practices vary widely across the colleges, but a statewide survey revealed that standardized placement exams remain central and frequently dominant in the placement process (Rodriguez, Mejia, & Johnson, 2016; WestEd, 2011). A closer look at the nine community colleges of the Los Angeles County Community College District confirmed the variability in practices observed across the state but also the centrality of standardized placement exams in the assessment of student readiness (Melguizo, Kosiewicz, Prather, & Bos, 2014; Ngo & Kwon, 2014).

There are few state guidelines to assist California's community colleges in implementing multiple measures assessment and placement processes, which perhaps is unsurprising in light of the evident variability in placement practices and the limited extent to which multiple measures placement has been implemented. Most of the existing requirements and guidelines concern the validation of placement tests (California Community College Chancellor's Office, 2001), while colleges currently are not required to validate their multiple measure policies (California Community College Chancellor's Office, 2011). Nevertheless, the Academic Senate for California Community Colleges (ASCCC) recently reaffirmed the importance of multiple measures placement practices (Academic Senate for California Community Colleges, 2013), and an ASCCC task force concluded that "the inclusion of multiple measures in our assessment processes is an important step toward the goal of improving the accuracy of assessments" (Grimes-Hillman, Holcroft, Fulks, Lee, & Smith, 2014, p. 7).

## DATA & MEASURES

### Data Source

Data for this project were drawn from Cal-PASS Plus (CPP), a project of the California Community Colleges Chancellor's Office (CCCCO). CPP is a voluntary, intersegmental data repository that includes educational and workplace outcome data on California residents. CPP was initiated in the mid-2000s, and many public K-12 schools in California, all of the state's community colleges, and some public four-year institutions now contribute data to CPP. Data for the community colleges were provided by the CCCCO and are available from Summer 1992 to nearly present day. Participation by K-12 schools initially was low but has grown over time. At the time that this study began, 1,882 California public high schools had submitted student transcript data to CPP, or about 46% of the 4,076 public high schools in the state.

**Sample**

The theoretical universe for this study is composed of all California community college students who enrolled in a developmental or college-level math or English course between Summer 1992 and Spring 2015. For the purposes of this study, we define developmental math as included arithmetic, pre-algebra, beginning algebra, and intermediate algebra, while college-level (transfer-level) math includes general education math, statistics, college algebra, trigonometry, pre-calculus, and calculus I. Like math, development English includes four levels of coursework, but we focus on only one category of college-level English, encompassing college composition and writing. We analyze each level of developmental math and English separately, just as we analyze each category of college-level math and English separately.

The primary limitation imposed on the analytical sample is that students have a valid academic year GPA for the 9th, 10th, 11th, and 12th grades. Because K-12 districts varied in the amount of historical data that they provided to CPP, the analytical sample is limited to 245,020 students, with most (96%) of these having entered the CCC system between Fall 2007 and Summer 2014.

Of this sample, 216,892 students enrolled in a developmental or college-level English course, and 201,986 students enrolled in a developmental or college-level math course, as defined here. Note that the sum of these figures is greater than the total number of students in the sample because some students enrolled in both math and English. Note also that we did not impose on the sample the requirement that these math/English enrollments occur in a student's first year of college. Nevertheless, 70 percent of the students in our sample enrolled in a developmental or college-level English course in their first year, as did 48 percent of students who enrolled in a developmental or college-level math course. Characteristics of the students in the sample are presented in Table 1.

[insert Table 1 about here]

**Measures**

The primary outcome of interest for each analytical sample was a dichotomous indicator of successful completion of a students' first math and/or first English course, focusing on the developmental and college-level math and English courses described previously. Successful completion is defined as a grade of A, B, C, Credit, or Pass, while a non-successful outcome is defined as a grade of D, F, No Credit, No Pass, or Withdrawal. Incomplete grades (e.g., IA, IB, IC, ID, IF) were recoded to their default value.

To test if results differ between students who transition directly from high school to college with no delay (i.e., direct matriculants, hereafter DM) and students who delay starting college by one or more years after the end of high school (non-direct matriculants, hereafter NDM), we created two sets of models. The DM models use data only through 11th grade because DM students are still enrolled in high school when placement decisions are made, making 12th grade academic performance unobserved with respect to placement decision-making by colleges. The NDM models use all data through 12th grade. The analytical samples used for the DM and NDM models were the same.

The independent variables used to predict successful completion of the student's first math course for the DM analyses included cumulative high school GPA (measured on a 4-point scale) at of the end of 11th grade, grade points achieved in math in each of 9th grade, 10th grade, and 11th grade (again measured on a 4-point scale), and individual dichotomous indicators of whether a student ever attempted and, if so, whether they achieved a grade of B or better, B-, C+, or C in each of the following courses: algebra 1, algebra 2, geometry, trigonometry, pre-calculus, calculus, statistics, or any Advanced Placement (AP) math.

For the NDM math analyses, the independent variables differed only in that 12th grade coursework and grade points also were included, and cumulative high school GPA as of the end of 12th grade was substituted for 11th grade cumulative GPA. Additionally, we included a measure of delay of math, which we operationalized as the number of primary terms (excluding summers) between completion of the last high school math course and enrollment in the first community college math course.

The independent variables for the DM English analyses were similar and included cumulative high school GPA as of the end of 11th grade, grade points achieved in English in each of 9th grade, 10th grade, and 11th grade, and dichotomous indicators of whether a student ever attempted and, if so, whether they achieved a grade of B (or better), B-, C+, or C in each of the following courses: English as a Second Language (ESL), any Expository English, a California State University (CSU) designed Expository Reading and Writing Course (ERWC), developmental English, or AP English. The independent variables for the NDM analyses differed in the substitution of cumulative GPA at the end of 12th grade for 11th grade cumulative GPA, the inclusion of grade points achieved in English in 12th grade, and a measure of delay between last high school English course and first community college English course.

## METHOD OF ANALYSIS

### Decision Trees

We use decision tree methods (Breiman, Friedman, Olshen, & Stone, 1984; Dietterich, 2000; Murthy, 1998; Quinlan, 1986) to analyze relationships between the measures of high school achievement and success in students' first math and English courses. Decision trees were developed as a machine learning and data mining tool but are finding use in education research as well (Ahmed & Elaraby, 2014; Willett, Hayward, & Dahlstrom, 2008). Decision trees draw on a set of user-specified variables to classify observations into an increasingly number of groups of progressively smaller size. At each step of the classification process, the algorithm selects the one variable and, further, the threshold of values on that variable that best sort the observations into two groups, seeking to maximize the homogeneity of a specified outcome within each group. Information measures like the Gini-Simpson index are commonly used to determine within-group homogeneity. The algorithm stops classifying observations when the tree reaches a specified complexity parameter, chosen to minimize cross-validation error.

Decision tree methods offer several advantages over traditional regression methods for this study. First, decision trees are relatively assumption-free; they do not have implicit distributional or relational assumptions, reducing challenges to the potential fit of the data or models. Second, decision tree methods are less restrictive in dealing with missing data than are regression methods (Ding & Simonoff, 2010). Third, decision trees support the analysis of nonlinear and non-additive relationships without variable transformations or the specification of interactions. Finally, decision trees categorize students into discrete groups according to their likelihood of a particular outcome using simple *if-then* logical statements. These statements can be presented visually in a manner that is intuitive and relatively easy to understand for lay audiences. This last aspect of decision tree methods is especially important for this study because it facilitates the communication of results to college administrators, educators, and practitioners. In turn, the clear communication helps to garner support for the changes to institutional placement practices that may result from bringing research findings to bear. All decision trees in this study were produced using R with the -*rpart*- package (Therneau, Atkinson, & Ripley, 2015).

We developed decision trees predicting success in each of nine levels of math coursework and four levels of English coursework, separately for direct matriculants (DMs) and non-direct matriculants (NDMs). We did not develop decision trees for the lowest level of math and English because these courses often are open-access without readiness requirements. We thus developed a total of 18 decision trees for math and 8 decision trees for English.

**Recursive Process**

Given that our outcomes were binary indicators of success in particular courses, each reflecting a level of skill in a curricular hierarchy, we developed a Poisson decision tree for each course using a *recursive* approach. In particular, we estimated models starting with the college-level courses in each subject, using subsamples of our larger analytical sample composed of all students whose first math/English course was one of these college-level courses. For example, the students whose first math course was calculus I (first-semester calculus) constituted the subsample used for the calculus I decision tree, while the students whose first math course was pre-calculus constituted the subsample used for the pre-calculus decision tree, and so on.

Then, once placement rules were determined for all college-level courses in a subject, students who met the placement rules for one or more of these college-level courses were excluded from the analytical subsample used to develop the decision trees for the highest developmental course in the subject, as well as the subsamples used for all lower-level courses.[1]

[1] As a point of clarification, we note that there are a large number of potential first college-level math courses in the CCC system, including calculus, pre-calculus, college algebra, trigonometry, statistics, and general education math. We did not apply our recursive approach *within* these six college-level math courses. The analytical subsample for each of these courses included all

This exclusion occurred regardless of the actual skill-level of a student's first math/English course. For example, a student whose first math course was intermediate algebra (the highest-skill developmental math course) was *included* in the analytical subsample used to develop the intermediate algebra decision tree only if the student's first math course was intermediate algebra *and* the student's high school achievement would not have qualified him/her to enroll in one or more of the college-level math courses, given the decision trees for the college-level math courses and their corresponding placement rules. This recursive process was continued for the next highest developmental course in the subject (e.g., beginning algebra in mathematics), with students meeting placement rules for the highest developmental course or one or more college-level courses being excluded from the analytical subsample.

Our recursive approach prevents the artificial inflation of expected success rates in the decision tree output, since students placing into higher-level courses are excluded from lower-level trees. Table 2 provides the number of students who attempted each math and English course as their first course in the subject, as well as the size of each analytical subsample resulting from our recursive approach.

[insert Table 2 about here]

---

students who attempted that course as their first math course, regardless of whether the decision tree and corresponding placement rules for another college-level math course would have qualified the students for placement in that other course. For example, students who would be predicted to be successful in pre-calculus based on their high school achievement, but who enrolled in college algebra as their first math course, were retained in the analytical subsample for college algebra. Once the decision tress for the six college-level math courses were finalized, we excluded from the analytical sample for the highest *developmental* math course all students who met the placement rules for at least one of the six *college-level* math courses. We then continued our recursive approach, as described, for the next highest developmental math course.

**Success Criteria**

The decision tree for each course (whether math or English) creates distinct groups of students, referred to as *nodes*, each of which has a predicted success rate — the proportion of students in a given group who are predicted to pass the focal course. Determining which nodes define the curricular placement rules for a course requires an a priori success criterion, constituting the minimum expected course success rate. Nodes with a predicted success rate at or above this criterion then define the placement rules for the course. In other words, the groups of students identified as eligible to be placed at a specific level of coursework would have a predicted success rate equal to or higher than the criterion.

Success criteria for this study were selected conservatively to reflect historic success rates in the courses, to ensure that students placed under rules developed from the decision trees were likely to succeed, and that the application of such placement rules would not negatively impact overall course success rates. Table 3 summarizes the success criteria selected for this study, as well as the actual (observed) rate of success in each course for students in the sample.

[insert Table 3 about here]

## FINDINGS

In the subsections that follow, we present interpretations of decision trees for two college-level math courses — statistics and college algebra — for direct matriculants (DMs) to illuminate the relationship between the results and the development of student placement rules. We then summarize the results for all decision trees for math, focusing on patterns evident across the courses and similarities between DMs and non-direct matriculants (NDMs). Finally, we present the interpretation of the decision tree for college-level English, and then summarize the results for all decision trees for English.

**Mathematics**

A notated decision tree for college-level statistics for DMs is presented in Figure 1. At the root node (node 1), the analytical sample is split by cumulative high school GPA as of the end of 11th grade. The right-hand branch includes students who have cumulative high school GPAs equal to or greater than 3.0, while the left-hand branch includes the remaining students with cumulative high school GPAs of less than 3.0. The appearance of cumulative GPA in the root node, at the highest level of the tree, indicates that it is a key predictor of success in college statistics among the variables included in our analysis.

[insert Figure 1 about here]

Reading down the right-hand side of the tree, the first internal node (node 2) splits the subsample of students with cumulative high school GPAs at or above 3.0 into a subgroup with GPAs at or above 3.3 (the right-hand branch) and a subgroup with GPAs between 3.0 and 3.2 (the left-hand branch). To the right is a terminal node (node 3); terminal nodes also are referred to as leaves. This leaf contains 30% of the overall DM sample for statistics and has a predicted probability of success in statistics of 0.88. One would interpret this leaf as indicating that 88% of community college DMs with cumulative high school GPAs of 3.3 or higher are predicted to succeed in college statistics as their first math course — a high probability of success. Of note, this leaf combined with the other leaves in the figure account for 100 percent of the sample that was used to generate this decision tree.

In the left-hand branch from node 2 are students with cumulative high school GPAs between 3.0 and 3.2. In node 4, these students are split by whether they enrolled in pre-calculus in high school. In our analysis, enrollment in pre-calculus is a simple dichotomous variable assigned a value of 1 if the student attempted pre-calculus prior to the $12^{th}$ grade and 0 otherwise.

Students who did so, and who had cumulative high school GPAs between 3.0 and 3.2, are found in terminal node 5. They compose 8% of the DM sample with a predicted 0.81 probability of success in statistics as a first math course. Those who did not take pre-calculus are found in terminal node 6, containing 16% of the sample with an estimated 0.70 probability of success in statistics as a first math course.

Returning to node 1, the left-hand branch contains students with cumulative high school GPAs of less than 3.0. Tracing this and subsequent branches to the terminal nodes, we find only one terminal node with a predicted probability of success of 0.70 or greater, which is node 9. This node contains students with cumulative high school GPAs between 2.3 and 2.9, who also completed high school pre-calculus with a grade of C or better.

As noted in Table 3, the success criterion for transfer-level courses is 0.70. Hence, a set of placement rules drawn from this decision tree indicates that DMs with cumulative high school GPAs of 3.0 or greater may be placed in statistics. In addition, DMs with cumulative high school GPAs between 2.3 and 2.9 may be placed in statistics if they also completed high school pre-calculus with a grade of C or better. Otherwise, the students should be placed in a lower-level math course.

The decision tree predicting success in college algebra for direct DMs is presented in Figure 2. As with statistics, cumulative high school GPA at the end of $11^{th}$ grade determines the split of the sample at the root node and at the two subsequent nodes. In the lower branches, we observe splits based on performance in high school pre-calculus, algebra 2, and algebra 1. For example, the right-most path leads to a terminal node containing the 7 percent of the sample who had cumulative high school GPAs of 3.2 or greater, and who also completed high school pre-calculus with a B- or better by the end of 11th grade. These students have a predicted 0.89 probability of success in college algebra as a first math course.

[insert Figure 2 about here]

The next terminal node to the left contains the 21 percent of the sample that met the same GPA threshold but did not complete high school pre-calculus with a B- or better. Their predicted probability of success is 0.79. One more terminal node to the left is the node containing the 5% of the sample with cumulative high school GPAs between 2.9 and 3.1, who also completed high school pre-calculus with a C or better. These students have a predicted probability of success of 0.75.

The three right-most terminal nodes all exceed the success criterion of 0.70 (Table 3), while none of the other terminal nodes do so. Thus, the placement rules for college algebra for DMs is a cumulative high school GPA of 3.2 or higher, or a GPA between 2.9 and 3.1 supplemented by a grade of C or better in high school pre-calculus by the end of $11^{th}$ grade. These placement rules are similar to the rules for statistics with the exception that the minimum GPA for successful completers of high school pre-calculus is somewhat lower for placement in statistics (GPA $\geq$ 2.3) than it is for placement in college algebra (GPA $\geq$ 2.9).

In Table 4, we summarize the placement rules derived from the decision trees for all levels of math and for both DMs and NDMs. Reviewing the results, it is clear that cumulative high school GPA is central to predicting the success of students, both those who enter college immediately after high school and those who do not. In addition, we note that the placement rules for DMs and NDMs are similar in most cases, indicating that the information about students' readiness that is added by their achievement in $12^{th}$ grade does not fundamentally alter predictions about their success.

[insert Table 4 about here]

**English**

The decision tree for success in transfer-level English for DMs is presented in Figure 3. As with the math trees, cumulative high school GPA again emerges as a key predictor of success, as evident at the root node and both subsequent nodes. The leaf on the far right of the figure indicates that students with cumulative high school GPAs at or above 3.1 have a 0.87 predicted probability of success in transfer-level English, while the next node to the left indicates that students with GPAs between 2.6 and 3.0 have a 0.73 predicted probability of success. Only these two nodes meet the 0.70 success criterion. Therefore, the placement rule derived from this decision tree is that DMs with cumulative high school GPAs of 2.6 or higher should be placed in transfer-level English.

[insert Figure 3 about here]

In Table 5, we observe that high school GPA is a consistently important predictor of success in all levels of English and for both DM and NDMs, as is the case with math. We note, though, that the minimum GPA to access college-level English is somewhat lower than the minimum GPA to access college-level math, which is consistent with prior work (Bahr, 2016). Also like math, placement rules for DM and NDMs are fairly similar.

[insert Table 5 about here]

## MODEL VALIDATION

There are two important questions to ask regarding the validity of these results. First, does assessing student readiness with multiple measures of high school achievement improve success in the target course, relative to assessing students with placement tests? Second, are the results replicable in other samples?

To answer these questions, we began by developing two separate testing samples. Concerning the first of these testing samples, because CPP is a voluntary and ongoing data sharing system, additional data became available while the work of this study was underway. These new data addressed the 2015-2016 academic year. Drawing on these data, we applied the same sample selection criteria of valid information on $9^{th}$, $10^{th}$, $11^{th}$, and $12^{th}$ grade GPA to create a new sample referred to here as *testing sample 1*.

Regarding the second testing sample, a number of students were excluded from our analytical sample or from testing sample 1 because they were missing valid information on high school GPA in one or more of the grade levels. However, some of these students had sufficient information on high school achievement to reach a placement recommendation, given the placement rules derived from our decision tree analyses. These students with partially missing but sufficient information for placement decisions were included in *testing sample 2*.

To validate our placement rules, we first applied the rules to the analytical sample and to the two testing samples, with the result that we were able to determine the highest placement level for each student in math and English. For each of the three samples, we calculated the success rate in the particular math or English course for students whose highest placement matched their actual first course in the subject. Finally, we compared these success rates with (1) the specified success criteria, and (2) the corresponding figures for students who had any high school information, disregarding the placement rules developed in this study. The results are presented in Table 6.

[insert Table 6 about here]

The results presented in Table 6 support the validity of the placement rules derived from the decision trees. In particular, we note that, with only one exception, the observed success rates for the analytical sample and two testing samples exceed those of students with any high school information, most of whom presumably were placed using the standardized placement tests historically favored in the colleges of the CCC system. The one exception is in developmental English three levels below college-level, where testing sample 1 has an observed success rate equal to that of students with any high school information.

In addition, success rates of the analytical sample and testing samples all exceed the success criteria, save for one. In testing sample 2, the 0.58 success rate in beginning algebra is slightly lower than the success criterion of 0.60.

Finally, success rates generally are very stable across the analytical sample and two testing samples. Only trigonometry varies meaningfully, with students in the analytical sample having a success rate of 0.81, while success rates of 0.66 and 0.74 are observed in testing samples 1 and 2, respectively.

## DISCUSSION & CONCLUSION

### Summary

The open admissions policies of community colleges guarantee access to postsecondary education but also ensure a student population of highly varied educational experiences and readiness for college-level coursework. In the interest of helping students to identify and remediate any skill shortfalls, community colleges employ various assessment tools and then sort students into coursework appropriate for their levels of readiness. Among these assessment tools, standardized placement exams have become a mainstay of community colleges. However, recent studies have shown that students' potential to succeed in math and English coursework often is underestimated by placement tests, with many students being placed into remedial courses despite being likely to succeed in college-level coursework (Bahr, 2016; Scott-Clayton et al., 2014b). In turn, being placed in remedial courses increases students' risk of leaving college without a postsecondary credential (Burdman, 2012). Thus, the means by which many community colleges strive to increase students' chances of success in college can prove ultimately to be an obstacle to their success.

From the perspective of community college instructors, however, the occurrence of student under-placement is largely invisible because the counterfactual of the same student taking a higher-level course is unobserved. More apparent to instructors is over-placement resulting in students struggling in a course, but evidence indicates that over-placement is much less common with placement tests than is under-placement (Scott-Clayton et al., 2014b). The difference in the visibility of these two phenomena — under-placement and over-placement — can lead to inaccurate assumptions about the effectiveness of institutional placement policies. For example, some community colleges adopt more stringent placement polices than do four-year institutions in equivalent courses (Fields & Parsad, 2012). More importantly, the relative invisibility of under-placement has resulted in the problem remaining largely unaddressed in most states.

The limited research to date has indicated that broad measures of high school achievement (e.g., GPA) can be used to generate placement recommendations that are more accurate, on average, than are placement recommendations based on placement tests, resulting in fewer students being under- or over-placed in math and English coursework (Bahr, 2016; Scott-Clayton et al., 2014b). This likely is a result of the fact that such measures are based on information collected over a lengthy period of time by many observers (instructors) in a variety of learning contexts. Moreover, they reflect many means by which students may demonstrate their academic capabilities (e.g., exams, papers, oral reports), and they also capture a range of behaviors and skills that are necessary for success in school (e.g., timeliness, organization, diligence, planning). In contrast, placement exams are executed in just minutes and are largely unidimensional, reflecting a single mode of academic capability.

The study described here was conducted as part of the California Multiple Measures Assessment Project (MMAP) and is one of the first *statewide* initiatives to examine empirically the use of high school achievement to place community college students in math and English coursework. Historically, community colleges in California have relied heavily on placement exams (Rodriguez et al., 2016; WestEd, 2011), despite system-level acknowledgement of the importance of using multiple measures of readiness (California Community College Chancellor's Office, 2011). We used decision tree methods to identify key measures of high school achievement that are strong predictors of success in community college math and English courses. We also determined the thresholds on these measures above which success in a given math or English course is likely. Importantly, our approach and the large size of our dataset allowed us to distinguish levels of development coursework and level of college coursework, rather than collapsing developmental and college-level coursework into monolithic categories, which has been an important limitation of prior research.

The results of our decision tree analyses indicate that cumulative high school GPA is the most consistently useful predictor of success across levels of math and English coursework. Further, the thresholds of high school achievement that indicate readiness for particular math and English courses are similar for students who matriculate directly from high school and students who delay matriculation. Drawing on predictions about students' likelihood of passing each math and English course, we developed placement rules sets, which we validated with two testing samples.

**Communicating Results to Stakeholders**

In communicating this research and the placement rule sets to faculty, administrators, and policymakers in California, an important strength of our decision tree methods has been the interpretability of the results. The connection between the analyses and the placement rules is transparent even to stakeholders who are not versant in quantitative research, which increases confidence in the empirical grounding of the rules.

At the same time, however, the prominence of high school GPA in the results has led some stakeholders to question whether the placement rule sets actually reflect the *multiple measures assessment* that the work was charged to develop. This is an understandable misinterpretation of the results because, unlike regression methods, decision trees do not display coefficients for all included variables regardless of their relative weight in predicting the outcome of interest. Consequently, it has been important to clearly articulate to stakeholders the large number of high school achievement measures that were included in the analyses and the logic by which information on one or two measures of high school achievement (e.g., GPA, highest math or English course) can render superfluous information on additional measures of high school achievement (on average).

Another challenge with acceptance of the rule sets has been concern by some colleges that using high school achievement may result in excessive rates of over-placement, rather than a reduction in the largely unobserved and unacknowledged problem of under-placement. Important in this regard has been sharing the results experienced by colleges that have piloted the placement rules (e.g. Cooper, 2016; Huang, Hsieh, & Lopez, 2016). Early findings indicate that using measures of high school achievement to place students increased the number of students in college-level coursework while still maintaining historic rates of success in these courses. In other words, students placed via placement tests did not differ in their rate of success from students placed via high school achievement, on average, but more students qualified to take college-level coursework based on high school achievement.

Furthermore, to ease acceptance of the math rule sets in particular, supplementary minimum high school course completion requirements were added to the placement rules. For placement in calculus or pre-calculus, new students were expected to meet the placement rules *and* have completed high school pre-calculus, trigonometry, or a higher-level course. For placement in trigonometry or college algebra, students must meet the placement rules and also have completed high school algebra II or a higher-level course. For placement in general education math, statistics, or intermediate algebra, the minimum is high school algebra I. Placement in beginning algebra, pre-algebra, or arithmetic depends on the placement rule sets only and requires no minimum level of high school coursework. Empirically speaking, these additional requirements have very little effect on the number of students placed in a given level of math because the large majority of students who met the placement rules for that level of math also met the additional high school coursework requirement. Still, the additional requirements helped to ease concerns about the possibility of student over-placement in math.

**Implications & Future Research**

Future research on the use of high school achievement to place community college students should estimate the effect of such policy changes on the rate of direct enrollment in college-level coursework and should disaggregate this effect by student sociodemographic characteristics, such as socioeconomic status, gender, and race/ethnicity. Unfortunately, a lack of consistent data on students' placement test scores prevented a statewide investigation of that sort in this study. Evidence from one prior study in an urban setting suggests that placing students based on high school achievement alters the racial/ethnic composition of college-level math and English courses, increasing the fraction of Hispanic students but decreasing the fraction of Black and Asian students (Scott-Clayton & Stacey, 2015). However, the authors' projections held constant the overall rate at which students are directed to college-level coursework, which would not be the case in reality if, as the evidence indicates, under-placement is more common than is over-placement. Instead, one would expect increased enrollment in college-level coursework across the spectrum of student backgrounds and perhaps most among students of lower socioeconomic status who may receive less counsel about the importance of preparing for placement tests.

The potential increase in the number of students placed in college-level coursework when using measures of high school achievement raises important considerations for institutions. On one hand, administrators and faculty will need to be prepared to accommodate greater demand for college-level coursework, likely from a student population that is more heterogeneous with respect to academic backgrounds. On the other hand, administrators and faculty will need to be prepared for declining enrollment in developmental courses as well as a possible homogenization of student skills in the developmental courses that remain, as students who previously would have been under-placed are directed upward to courses that more accurately match their readiness. These changes will have consequences both for the faculty who teach college-level and developmental math and English courses and for the nature of instruction in these courses, and these consequences should be investigated in future research.

Another important implication of substituting measures of high school achievement for college-administered placement tests is that colleges will be incentivized to work more closely with local high schools to prepare students for college, rather than simply critiquing the preparation provided by high schools. Future research should investigate how this change alters the dialogue between community colleges and local high schools, as well as the discourse about local high schools among college faculty and administrators.

In addition to aforementioned lines of inquiry, three aspects of our work require further investigation in future research. First, we drew a distinction in our analyses between the information that would be available to colleges about students who matriculate directly from high school and those who delay matriculation. However, more research is needed on the relationship between the length of the delay between high school graduation and college enrollment and the extent to which measures of high school achievement can be used to predict success in math and English coursework. One might ask whether the utility of high school achievement measures for predicting success fades over time, remains relatively constant, or possibly grows.

Second, as noted, we investigated *cumulative* GPA as reported by high schools and found it to be a key predictor of success in math and English coursework. It may be worthwhile, though, for future research to consider more nuanced variations of GPA such as subject-specific GPA (e.g., cumulative GPA in math courses, cumulative GPA in English courses).

Finally, this study benefited from access to information about students that was reported directly by high schools. Yet, many community colleges and system offices currently do not have access to high school transcripts. Therefore, it will be important for future research to investigate the viability of self-reported information about high school achievement in place of information reported directly by high schools.

**REFERENCES**

Academic Senate for California Community Colleges. (2013). Use of Multiple Measures in Common Statewide Assessment Exams (Resolution Number 18.01). Retrieved May 6, 2016, from http://www.asccc.org/resolutions/use-multiple-measures-common-statewide-assessment-exams

Ahmed, A. B., & Elaraby, I. S. (2014). Data mining: A prediction for student's performance using classification method. *World Journal of Computer Application and Technology*, *2*(2), 43–47.

Attewell, P., Lavin, D., Domina, T., & Levey, T. (2006). New evidence on college remediation. *Journal of Higher Education*, 886–924.

Bahr, P. R. (2008). Does mathematics remediation work?: A comparative analysis of academic attainment among community college students. *Research in Higher Education*, *49*, 420–450.

Bahr, P. R. (2012). Deconstructing remediation in the community college: Exploring associations between course-taking patterns, course outcomes, and attrition from the remedial math and remedial writing sequences. *Research in Higher Education*, *53*, 661–693.

Bahr, P. R. (2016). *Replacing placement tests in Michigan's community colleges*. Ann Arbor, Michigan: Center for the Study of Higher and Postsecondary Education, University of Michigan.

Bahr, P. R., & Gross, J. (2016). Community colleges. In M. N. Bastedo, P. G. Altbach, & P. J. Gumport (Eds.), *American higher education in the 21st century, 4th edition* (pp. 462–502). Baltimore, Maryland: John Hopkins University.

Bahr, P. R., Jackson, G., McNaughtan, J., Oster, M., & Gross, J. (2017). Unrealized potential: Community college pathways to STEM baccalaureate degrees. *Journal of Higher Education*, *8*(3), 430–478.

Bailey, T. (2009). Challenge and opportunity: Rethinking the role and function of developmental education in community college. *New Directions for Community Colleges*, *2009*(145), 11–30.

Bailey, T., Bashford, J., Boatman, A., Squires, J., Weiss, M., Doyle, W., … Young, S. H. (2016). *Strategies for postsecondary students in developmental education–A practice guide for college and university administrators, advisors, and faculty*. Washington, DC: Institute of Education Sciences, What Works Clearinghouse. Retrieved from https://ies.ed.gov/ncee/wwc/PracticeGuide/23

Bailey, T., Jeong, D. W., & Cho, S. (2010). Referral, enrollment, and completion in developmental education sequences in community colleges. *Economics of Education Review*, *29*(2), 255–270.

Belfield, C., & Crosta, P. M. (2012). *Predicting success in college: The importance of placement tests and high school transcripts (CCRC Working Paper No. 42)* (Working Paper). New York: Columbia University, Teachers College, Community College Research Center.

Bettinger, E., & Long, B. T. (2005). Remediation at the community college: Student participation and outcomes. *New Directions for Community Colleges*, *129*, 17–26.

Bickerstaff, S., Fay, M. P., & Trimble, M. J. (2016). *Modularization in developmental mathematics in two states: Implementation and early outcomes*. New York, NY: Community College Research Center, Teachers College, Columbia University. Retrieved from https://ccrc.tc.columbia.edu/media/k2/attachments/modularization-developmental-mathematics-two-states.pdf

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Belmont, CA: Wadsworth.

Burdman, P. (2012). *Where to begin? The evolving role of placement exams for students starting college*. Washington, DC: Jobs for the Future. Retrieved from http://www.achievingthedream.org/sites/default/files/resources/Where_to_Begin.pdf

California Community College Chancellor's Office. (2001). *Standards, policies and procedures for the evaluation of assessment instruments used in the California community colleges, 4th Edition*. Sacramento, CA: California Community College Chancellor's Office. Retrieved from http://extranet.cccco.edu/Portals/1/SSSP/Matriculation/Assessment/ApprovedGuidelinesMarch2001.pdf

California Community College Chancellor's Office. (2011). *California Community College Matriculation Handbook*. Sacramento, CA: California Community College Chancellor's Office. Retrieved from http://extranet.cccco.edu/Portals/1/SSSP/Matriculation/MatriculationHandbookRevSeptember2011.pdf

California Community College Chancellor's Office. (2017). Policy in Action: Key Facts. Retrieved from http://californiacommunitycolleges.cccco.edu/PolicyInAction/KeyFacts.aspx

Cho, S.-W., Kopko, E., Jenkins, D., & Jaggars, S. S. (2012). *New evidence of success for community college remedial English students: Tracking the outcomes of students in the*

*Accelerated Learning Program (ALP).* (CCRC Working Paper No. 53). New York, NY: Community College Research Center, Columbia University.

Cooper, E. (2016). Soup to nuts: The full course on Sierra College's transition to multiple measures. Presented at the Annual conference of the Research and Planning Group for California Community Colleges, San Diego, CA.

Dietterich, T. G. (2000). An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Machine Learning*, *40*(2), 139–157.

Ding, Y., & Simonoff, J. (2010). An investigation of missing data methods for classification trees applied to binary response data. *Journal of Machine Learning Research*, *11*, 131–170.

Edgecombe, N., Jaggars, S. S., Xu, D., & Barragan, M. (2014). *Accelerating the integrated instruction of developmental reading and writing at Chabot College*. New York, NY: Community College Research Center, Teachers College, Columbia University. Retrieved from http://ccrc. tc.columbia.edu/publications/accelerating-integrated-instruction-at-chabot.html

Education Commission of the States. (2015). *Moving from single to multiple measures for college course placement*. Retrieved from http://www.ecs.org/moving-from-single-to-multiple-measures-for-college-course-placement/

Fay, M. P., Bickerstaff, S., & Hodara, M. (2013). *Why students do not prepare for math placement exams: Student perspectives.* (CCRC Research Brief No. 57). New York, NY: Community College Research Center, Teachers College, Columbia University. Retrieved from http://eric.ed.gov/?id=ED565657

Fields, R., & Parsad, B. (2012). *Tests and cut scores used for student placement in postsecondary education: Fall 2011*. Washington, DC: National Assessment Governing Board. Retrieved from https://www.nagb.org/content/nagb/assets/documents/commission/researchandresources/test-and-cut-scores-used-for-student-placement-in-postsecondary-education-fall-2011.pdf

Geiser, S., & Santelices, M. V. (2007). *Validity of high-school grades in predicting student success beyond the freshman year: High-school based record vs. standardized tests as indicators of four-year college outcomes*. Berkeley, CA: Center for Studies in Higher Education.

Grimes-Hillman, M., Holcroft, C., Fulks, J., Lee, D., & Smith, P. (2014). *The requirements and challenges of multiple measures in California Community Colleges.* Sacramento, CA: Academic Senate for California Community Colleges. Retrieved from http://www.asccc.org/sites/default/files/Multiple%20Measures%20in%20Assessment_0.pdf

Hayward, C., & Willett, T. (2014). *Curricular Redesign and Gatekeeper Completion: A Multi-College Evaluation of the California Acceleration Project*. Berkeley, CA: The Research and Planning Group for California Community Colleges.

Hiss, W. C., & Franks, V. W. (2014). *Defining promise: Optional standardized testing policies in American college and university admissions*. Arlington, VA: National Association for College Admission Counseling. Retrieved from https://www.luminafoundation.org/resources/defining-promise

Hodara, M., Jaggars, S. S., & Karp, M. M. (2012). *Improving developmental education assessment and placement: Lessons from community colleges across the country.* (CCRC

Working Paper No. 51). New York, NY: Community College Research Center, Teachers College, Columbia University.

Huang, T., Hsieh, C., & Lopez, K. (2016). Implementing Multiple Measures Assessment Project. Presented at the Annual conference of the Research and Planning Group for California Community Colleges, San Diego, CA.

Hughes, J., & Petscher, Y. (2016). *A Guide to Developing and Evaluating a College Readiness Screener*. (No. REL 2016-169). Washington, DC: Regional Educational Laboratory Southeast. Retrieved from http://ies.ed.gov/ncee/edlabs

Hughes, K. L., & Scott-Clayton, J. (2011). Assessing developmental assessment in community colleges. *Community College Review*, *39*(4), 327–351.

Jaggars, S. S., Hodara, M., & Stacey, G. W. (2013). *Designing Meaningful Developmental Reform*. New York, NY: Community College Research Center, Columbia University.

Jenkins, D., Speroni, C., Belfield, C., Jaggars, S. S., & Edgecombe, N. (2010). *A model for accelerating academic success of community college remedial English students: Is the Accelerated Learning Program (ALP) effective and affordable?* (CCRC Working Paper No. 21). New York, NY: Community College Research Center, Teachers College, Columbia University. Retrieved from http://ccrc.tc.columbia.edu/media/k2/attachments/remedial-english-alp-effective-affordable.pdf

Juszkiewicz, J. (2014). *Community college students and federal financial aid: A primer*. Washington, DC: American Association of Community Colleges. Retrieved from http://www.aacc.nche.edu/Publications/Reports/Documents/CCStudents_A_Primer.pdf

Kalamkarian, H. S., Raufman, J., & Edgecombe, N. (2015). *Statewide Developmental Education Reform: Early Implementation in Virginia and North Carolina*. New York, NY: Community College Research Center, Teachers College, Columbia University. Retrieved from http://ccrc.tc.columbia.edu/media/k2/attachments/statewide-developmental-education-reform-early-implementation.pdf

Mejia, M. C., Rodriguez, O., & Johnson, H. (2016). *Preparing students for success in California's community colleges*. San Francisco, CA: Public Policy Institute of California. Retrieved from http://www.ppic.org/main/publication.asp?i=1215

Melguizo, T., Kosiewicz, H., Prather, G., & Bos, J. (2014). How are community college students assessed and placed in developmental math? Grounding our understanding in reality. *The Journal of Higher Education*, *85*(5), 691–722.

Moss, B. G., Bahr, P. R., Arsenault, L., & Oster, M. (2016). *Improving community college students' performance on high-stakes placement tests: A randomized controlled trial of supplemental notification letters.* Ann Arbor, MI: Center for the Study of Higher and Postsecondary Education, University of Michigan. Retrieved from http://www.soe.umich.edu/people/profile/peter_riley_bahr/

Murthy, S. K. (1998). Automatic construction of decision trees from data: A multi-disciplinary survey. *Data Mining and Knowledge Discovery*, *2*(4), 345–389.

Ngo, F., & Kwon, W. W. (2014). Using multiple measures to make math placement decisions: implications for access and success in community colleges. *Research in Higher Education*, *56*, 5.

Ngo, F., & Kwon, W. W. (2015). Using multiple measures to make math placement decisions: Implications for access and success in community colleges. *Research in Higher Education*, *56*, 442–470.

Nodine, T., Dadgar, M., Venezia, A., & Bracco, K. R. (2013). *Acceleration in developmental education*. San Francisco, CA: WestEd. Retrieved from https://www.wested.org/online_pubs/resource1281.pdf

Parker, T. L., Bustillos, L. T., & Behringer, L. B. (2010). *Remedial and developmental education policy at a crossroads*. Boston, MA: Policy Research on Preparation, Access and Remedial Education (PRePARE), University of Massachusetts. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.188.7104&rep=rep1&type=pdf

Pearson. (2014). *Remedial education reform: Fixing the gateway to student success, college completion and employability*. Pearson. Retrieved from http://www.pearsoned.com/wp-content/uploads/2014/02/CCR-Briefing-Paper.pdf

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*, 81–106.

Rodriguez, O., Mejia, M. C., & Johnson, H. (2016). *Determining college readiness in California's community colleges: A survey of assessment and placement policies*. Sacramento, CA: Public Policy Institute of California. Retrieved from http://www.ppic.org/main/publication.asp?i=1216

Scott-Clayton, J. (2012a). *Do high-stakes placement exams predict college success?* New York, NY: Community College Research Center, Teachers College, Columbia University.

Scott-Clayton, J. (2012b). *Do high-stakes placement exams predict college success?* (CCRC Working Paper No. 41). New York, NY: Community College Research Center, Teachers

College, Columbia University. Retrieved from http://67.205.94.182/media/k2/attachments/high-stakes-predict-success.pdf

Scott-Clayton, J., Crosta, P. M., & Belfield, C. R. (2014a). Improving the Targeting of Treatment Evidence From College Remediation. *Educational Evaluation and Policy Analysis*, *36*(3), 371–393.

Scott-Clayton, J., Crosta, P. M., & Belfield, C. R. (2014b). Improving the targeting of treatment: Evidence from college remediation. *Educational Evaluation and Policy Analysis*, *36*(3), 371–393.

Scott-Clayton, J., & Stacey, G. W. (2015). *Improving the accuracy of remedial placement*. New York, NY: Community College Research Center, Teachers College, Columbia University. Retrieved from http://ccrc.tc.columbia.edu/media/k2/attachments/improving-accuracy-remedial-placement.pdf

Sedlacek, W. E. (2004). *Beyond the big test: Noncognitive assessment in higher education.* San Francisco, CA: Jossey-Bass. Retrieved from https://eric.ed.gov/?id=ED498559

Shulock, N., & Moore, C. (2007). *The rules of the game: How state policy creates barriers to degree completion and impedes student success in the California community colleges*. Sacramento, CA: Institute for Higher Education Leadership and Policy. Retrieved from http://edinsightscenter.org/Publications/ctl/ArticleView/mid/421/articleId/2061/categoryId/21/Rules-of-the-Game-How-State-Policy-Creates-Barriers-to-Degree-Completion-and-Impedes-Student-Success-in-the-California-Community-Colleges

Therneau, T., Atkinson, B., & Ripley, B. (2015). *Recursive partitioning and regression trees*. Retrieved from https://cran.r-project.org/web/packages/rpart/rpart.pdf

WestEd, R. W. for. (2011). *Types of multiple measures used in California community college mathematics, English, and English as a second language course placement: Summary report of survey results*. San Francisco, CA: WestEd.

Willett, T., Hayward, C., & Dahlstrom, E. (2008). *An early alert system for remediation needs of entering community college students: Leveraging the California Standards Test* (No. 2007036). Encinitas, CA: California Partnership for Achieving Student Success.

Table 1: Frequency distributions for selected variables for the overall sample (n = 245,020, except where noted)

| | | Frequency | Percentage |
|---|---|---|---|
| Gender | | | |
| | Male | 117,424 | 48% |
| | Female | 127,596 | 52% |
| Ethnicity | | | |
| | Asian | 23,316 | 10% |
| | Black | 14,455 | 6% |
| | Hispanic | 104,921 | 43% |
| | White | 67,628 | 28% |
| | Other | 8,117 | 3% |
| | Missing | 26,583 | 11% |
| Educational Goal in College | | | |
| | Transfer to four-year | 116,390 | 48% |
| | Terminal credential | 10,005 | 4% |
| | Other goal | 42,697 | 17% |
| | Undecided | 39,192 | 16% |
| | Not reported | 36,736 | 15% |
| | Percentage of students receiving fee waiver | 245,020 | 58% |
| | Percentage of students receiving Pell grant | 245,020 | 38% |
| | Percentage of students participating in EOPS | 245,020 | 8% |
| | Percentage of students participating in DSPS | 245,020 | 4% |
| Academic year of 12th grade | | | |
| | < 2006/07 | 10,121 | 4% |
| | 2006/07 | 28,983 | 12% |
| | 2007/08 | 44,089 | 18% |
| | 2008/09 | 36,760 | 15% |
| | 2009/10 | 52,169 | 21% |
| | 2010/11 | 33,641 | 14% |
| | 2011/12 | 27,714 | 11% |
| | > 2011/12 | 11,543 | 5% |
| 12th grade cumulative GPA | | | |
| | < 0.5 | 82 | 0% |
| | 0.5-0.9 | 903 | 0% |
| | 1.0-1.4 | 7,570 | 3% |
| | 1.5-1.9 | 32,689 | 13% |
| | 2.0-2.4 | 65,349 | 27% |
| | 2.5-2.9 | 72,258 | 29% |
| | 3.0-3.4 | 48,527 | 20% |
| | 3.5-4.0 | 17,642 | 7% |

| | | |
|---|---|---|
| Academic year first entered CCC system | | |
| < 2007/08 | 9,506 | 4% |
| 2007/08 | 20,116 | 8% |
| 2008/09 | 33,835 | 14% |
| 2009/10 | 36,710 | 15% |
| 2010/11 | 44,366 | 18% |
| 2011/12 | 37,752 | 15% |
| 2012/13 | 35,727 | 15% |
| > 2012/13 | 27,008 | 11% |
| First English course in college | | |
| Transfer level | 83,285 | 34% |
| One level below transfer | 74,040 | 30% |
| Two levels below transfer | 40,905 | 17% |
| Three levels below transfer | 16,135 | 7% |
| Four levels below transfer | 2,527 | 1% |
| Did not attempt English | 28,128 | 11% |
| First math course in college | | |
| Transfer level | 52,748 | 22% |
| One level below transfer | 58,446 | 24% |
| Two levels below transfer | 52,280 | 21% |
| Three levels below transfer | 28,230 | 12% |
| Four levels below transfer | 10,282 | 4% |
| Did not attempt math | 43,034 | 18% |
| Year of college in which first English course was taken | | |
| First | 151,655 | 62% |
| Second | 39,931 | 16% |
| Third | 18,282 | 7% |
| Fourth | 6,095 | 2% |
| Fifth or later | 929 | <1% |
| Did not attempt English | 28,128 | 11% |
| Year of college in which first math course was taken | | |
| First | 97,501 | 40% |
| Second | 54,021 | 22% |
| Third | 32,074 | 13% |
| Fourth | 13,795 | 6% |
| Fifth or later | 4,595 | 2% |
| Did not attempt math | 43,034 | 18% |
| Success on first attempt of first English course in college | 216,892 | 69% |
| Success on first attempt of first math course in college | 201,986 | 57% |

Note: DSPS refers to Disability Students Programs and Services. EOPS is Extended Opportunity Program and Services, state-funded program for students who are financially and educationally disadvantaged.

Table 2: Analytical sample sizes

| | | Number of Colleges | Total Number of Students | Recursive Analytical Sample |
|---|---|---|---|---|
| **English** | Transfer Level | 111 | 83,285 | N/A |
| | 1 Level Below Transfer | 110 | 74,040 | 37,924 |
| | 2 Levels Below Transfer | 105 | 40,905 | 16,729 |
| | 3 Levels Below Transfer | 77 | 16,135 | 3,142 |
| | 4 Levels Below Transfer | 20 | 2,527 | N/A |
| **Math** | Calculus | 103 | 8,833 | N/A |
| | Pre-Calculus | 93 | 9,446 | N/A |
| | Trigonometry | 93 | 4,074 | N/A |
| | College Algebra | 74 | 8,197 | N/A |
| | GE Math | 91 | 4,454 | N/A |
| | Statistics | 106 | 17,744 | N/A |
| | Algebra II | 111 | 58,446 | 30,099 |
| | Algebra I | 110 | 52,280 | 26,685 |
| | Pre-Algebra | 105 | 28,230 | 10,100 |
| | Arithmetic | 74 | 10,282 | N/A |

NOTE: "N/A" indicates that all students taking this course were included in the decision tree analysis for this level of math or English.

Table 3:  Success criteria for both math and English

| | Transfer-Level | One Level Below Transfer | Two Levels Below Transfer | Three Levels Below Transfer |
|---|---|---|---|---|
| Success criterion | 0.700 | 0.650 | 0.600 | 0.550 |
| Observed success rate | 0.676 | 0.614 | 0.565 | 0.567 |

NOTE:  The observed success rate is specific to the analytical sample.

Table 4: Placement rules for math, derived from decision trees

| College Math Course | Placement Rules for Direct Matriculants | Placement Rules for Non-Direct Matriculants |
|---|---|---|
| Calculus I | HS GPA ≥ 3.6<br>HS GPA ≥ 3.2 and completed HS pre-calculus ≥ C | HS GPA ≥ 3.5<br>HS GPA ≥ 3.1 and attempted HS calculus |
| Pre-calculus | HS GPA ≥ 3.4<br>HS GPA ≥ 2.6 and attempted HS calculus | HS GPA ≥ 3.3<br>HS GPA ≥ 3.0 and completed HS calculus ≥ C |
| Trigonometry | HS GPA ≥ 3.4<br>HS GPA ≥ 3.0 and completed HS pre-calculus ≥ C+<br>HS GPA ≥ 3.0 and completed HS algebra II ≥ B | HS GPA ≥ 3.3<br>HS GPA ≥ 2.8 and completed HS pre-calculus ≥ C |
| College algebra | HS GPA ≥ 3.2<br>HS GPA ≥2.9 and completed HS pre-calculus ≥ C | HS GPA ≥ 3.2<br>HS GPA ≥ 3.0 and completed HS pre-calculus ≥ C<br>HS GPA ≥ 3.0 and completed HS statistics ≥ C |
| General education math | HS GPA ≥ 3.3 | HS GPA ≥ 3.2<br>HS GPA ≥ 2.9 and completed HS statistics ≥ C |
| Statistics | HS GPA ≥ 3.0<br>HS GPA ≥ 2.3 and completed HS pre-calculus ≥ C | HS GPA ≥ 3.0<br>HS GPA ≥ 2.6 and completed HS pre-calculus ≥ C |
| Intermediate algebra | HS GPA ≥ 2.8 | HS GPA ≥ 2.9<br>HS GPA ≥ 2.5 and completed HS pre-calculus ≥ C |
| Beginning algebra | HS GPA ≥ 2.4 | HS GPA ≥ 2.5<br>HS GPA ≥ 2.3 and completed HS algebra II ≥ C |
| Pre-algebra | HS GPA ≥ 2.0 | HS GPA ≥ 2.1 |
| Arithmetic | All others students | All other students |

Table 5: Placement rules for English, derived from decision trees

| College English Course | Placement Rules for Direct Matriculants | Placement Rules for Non-Direct Matriculants |
|---|---|---|
| Transfer-level | HS GPA ≥ 2.6 | HS GPA ≥ 2.6 |
| One level below transfer | HS GPA ≥ 2.3 | HS GPA ≥ 2.2 and completed 12$^{th}$ grade English ≥ C |
| Two levels below transfer | HS GPA ≥ 2.0 | HS GPA ≥ 1.8 and completed 12$^{th}$ grade English ≥ D |
| Three levels below transfer | HS GPA ≥ 1.4 | HS GPA ≥ 1.7 |
| Four levels below transfer | All others students | All other students |

Table 6: Validation of placement rule sets, comparing course success rates for first math and English course enrollments

| | Success Criterion | Analytical Sample | Testing Sample 1 | Testing Sample 2 | Combined Testing Sample | Students with Any High School Information |
|---|---|---|---|---|---|---|
| **Math** | | | | | | |
| Calculus I | 0.70 | 0.79 | 0.77 | 0.81 | 0.80 | 0.65 |
| Pre-calculus | 0.70 | 0.74 | 0.74 | 0.74 | 0.74 | 0.60 |
| Trigonometry | 0.70 | 0.81 | 0.66 | 0.74 | 0.72 | 0.58 |
| College Algebra | 0.70 | 0.79 | 0.76 | 0.78 | 0.77 | 0.60 |
| General Education Math | 0.70 | 0.79 | 0.76 | 0.79 | 0.78 | 0.68 |
| Statistics | 0.70 | 0.78 | 0.78 | 0.79 | 0.79 | 0.68 |
| Intermediate Algebra | 0.65 | 0.70 | 0.71 | 0.70 | 0.70 | 0.55 |
| Beginning Algebra | 0.60 | 0.61 | 0.64 | 0.58 | 0.59 | 0.50 |
| Pre-algebra | 0.55 | 0.60 | 0.64 | 0.59 | 0.60 | 0.54 |
| Arithmetic | ---------- | 0.53 | 0.55 | 0.50 | 0.50 | 0.51 |
| *N* | | 201,986 | 27,624 | 237,115 | 264,739 | 1,224,903 |
| **English** | | | | | | |
| Transfer-level English | 0.70 | 0.78 | 0.77 | 0.77 | 0.77 | 0.70 |
| One level below transfer | 0.65 | 0.75 | 0.76 | 0.74 | 0.74 | 0.67 |
| Two levels below transfer | 0.60 | 0.70 | 0.67 | 0.69 | 0.69 | 0.64 |
| Three levels below transfer | 0.55 | 0.67 | 0.63 | 0.65 | 0.65 | 0.63 |
| Four levels below transfer | ---------- | 0.63 | 0.65 | 0.61 | 0.61 | 0.62 |
| *N* | | 216,892 | 31,041 | 320,842 | 351,883 | 1,280,406 |

Figure 1: Decision tree for statistics for direct matriculants

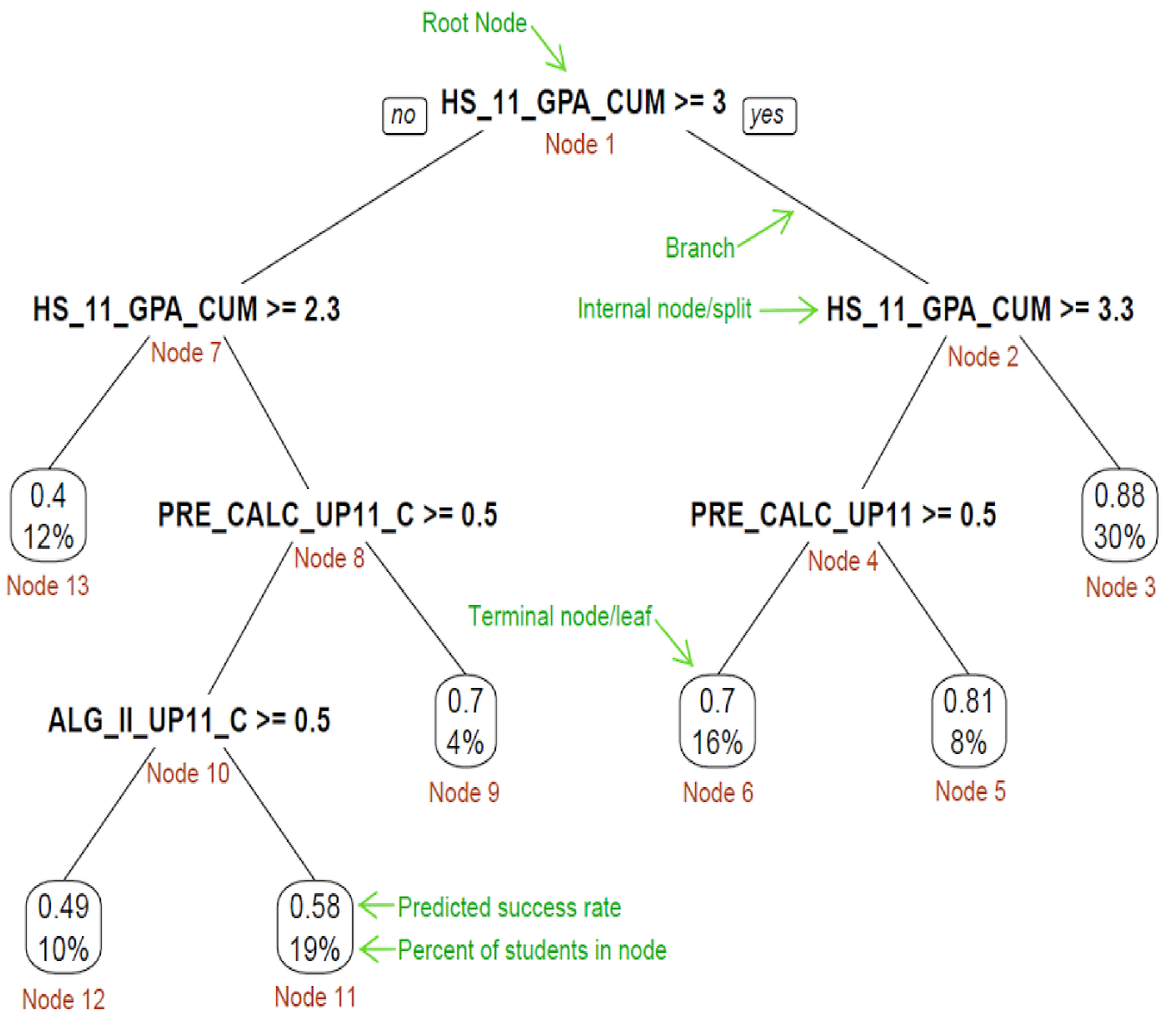

Figure 2: Decision tree for college algebra for direct matriculants

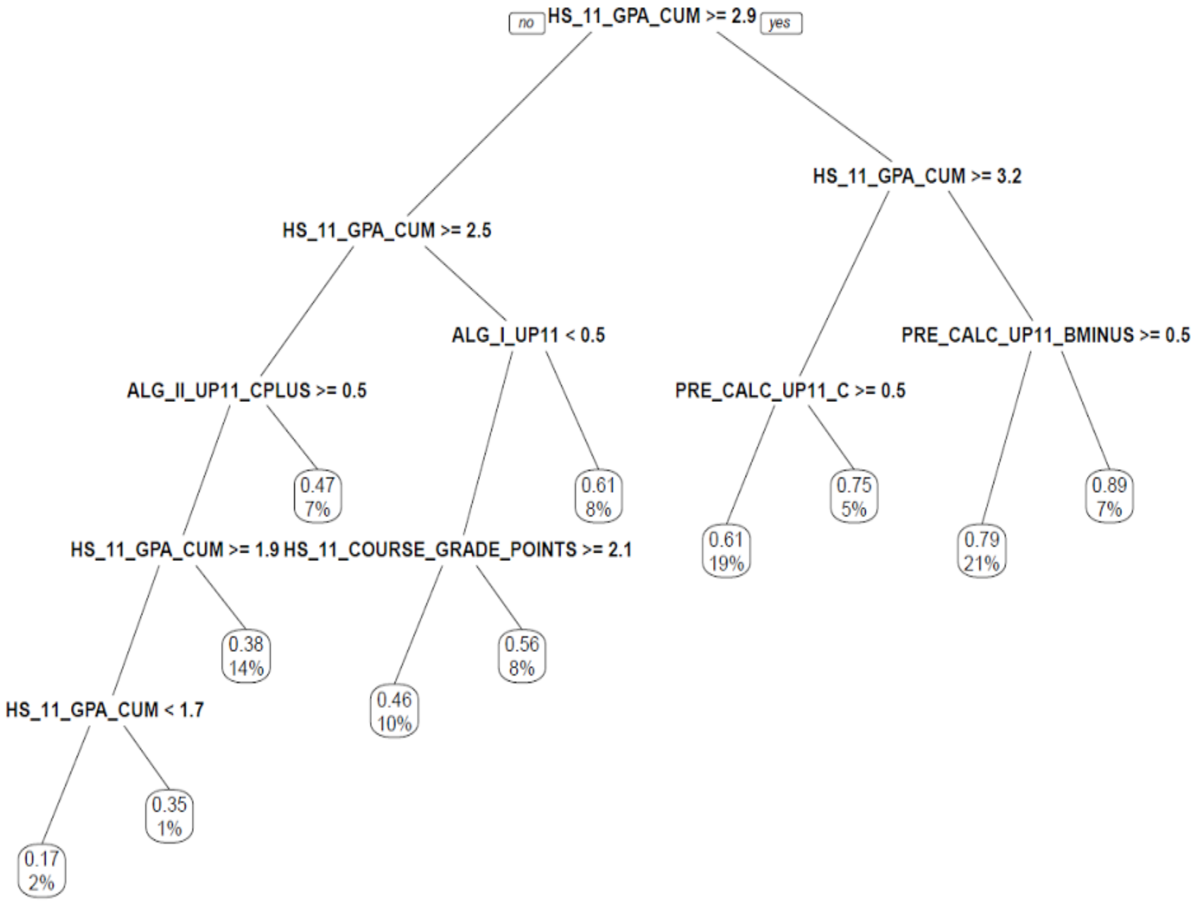

Figure 3: Decision tree for <u>transfer-level English</u> for direct matriculants