CCRC COMMUNITY COLLEGE RESEARCH CENTER

TEACHERS COLLEGE, COLUMBIA UNIVERSITY

Using Data Mining to Explore Why Community College Transfer Students Earn Bachelor's Degrees With Excess Credits

John Fink
Davis Jenkins
Elizabeth Kopko
Florence Xiaotao Ran

February 2018

CCRC Working Paper No. 100

Address correspondence to:

John Fink
Research Associate, Community College Research Center
Teachers College, Columbia University
525 W. 120th St., Box 174
New York, NY 10027
212-678-8187
Email: john.fink@tc.columbia.edu

Funding for this study was provided by the Bill & Melinda Gates Foundation. The authors of this report gratefully acknowledge our partners in the two anonymous states for providing the data for this study. The authors are also grateful for discussant comments from David B. Monaghan on an earlier version of this paper presented at the 2016 Association for Public Policy Analysis and Management Fall Research Conference.

Abstract

Community college transfer students encounter challenges progressing toward a bachelor's degree, leading to widespread transfer credit loss. This in turn may lower students' chances of credential completion and increase the time and costs for students, their families, and taxpayers. In this study we review three definitions of credit transfer inefficiency—*credit transferability*, *credit applicability*, and *excess credits among completers*—focusing on the last to examine why students who start at a community college and transfer to a four-year institution so often end up with excess credits that do not count toward a bachelor's degree. To shed light on credit transfer inefficiency, we examine the course-taking behaviors of community college transfer students who earn bachelor's degrees with numerous excess credits compared with transfer students who earn bachelor's degrees with few excess credits. We employ data-mining techniques to analyze student transcripts from two state systems, enabling us to examine a large number of variables that could explain the variation in students' excess credits at graduation. These variables include not only student demographics but also the types and timing of courses taken. Overall, we find more excess credits associated with several factors, including taking larger proportions of 100- and 200-level courses and smaller proportions of 300-level courses throughout students' progression toward completion, and taking 100-level courses in any subject—and specifically 100-level math courses—immediately after transferring to a four-year institution. Findings suggest that institutions could help students reduce credit transfer inefficiency by encouraging them to explore and choose a bachelor's degree major early on so they can take the required lower division (100- and 200-level) courses at the community college, thereby enabling them to take mostly upper division 300- and 400-level courses in their desired major field once they transfer to a four-year institution.

Table of Contents

1. Introduction

Community colleges are the entry point for many students who aspire to attain a four-year degree, but few of these students transfer, and even fewer earn bachelor's degrees. While over 80 percent of entering community college students intend to attain a bachelor's degree or higher, only one third transfer to a four-year institution within six years, and only one in seven earn a bachelor's degree (Jenkins & Fink, 2016). One of the major impediments these students encounter is the inefficiency of the credit transfer process. In a study based on a nationally representative sample of college students, Monaghan and Attewell (2015) found that only about 60 percent of community college entrants who transferred to a four-year institution were able to transfer the majority of their two-year college credits, and about 15 percent were hardly able to transfer any of their credits. The United States Government Accountability Office recently found that students who start at a community college lose many credits in the transfer process, and that even those students who transfer to a public four-year university are unable to apply 20 percent of the credits they earned at a community college toward a bachelor's degree (U.S. Government Accountability Office, 2017, Fig. 3). The loss of credits upon transfer not only decreases students' chances of completing a bachelor's degree (Monaghan & Attewell, 2015) but also increases the time and money they need to spend in order to earn the degree (Belfield, Fink, & Jenkins, 2017; Cullinane, 2014; Xu, Jaggars, & Fletcher, 2016).

In the current study, we examine why students who start in a community college and transfer to a four-year institution so often end up with excess credits that do not count toward a bachelor's degree. We investigate course-taking behaviors among community college transfer students who earn bachelor's degrees with numerous excess credits compared with transfer students who earn bachelor's degrees with few excess credits. To do this we use data-mining techniques to analyze longitudinal unit record data on students in two states who started at a community college and subsequently transferred and earned a bachelor's degree within the state college system. For comparison, we also conduct this analysis on a sample of students who started at a four-year college and completed a bachelor's degree. We examine a large number of variables that might affect the probability that students would earn excess credits, including not only student demographics but also the types of courses taken and the timing of those courses. We then examine the extent to which different sets of indicators surfaced through our data-mining analysis explain the variance in students' excess credits.

1.1 Measuring Credit Transfer Efficiency

Researchers have examined the efficiency of credit transfer using three measures: credit transferability, applicability of transfer credit, and excess credits among completers. *Credit transferability* refers to the number of credits students earned at one college that are accepted (or not) at another college. National and state studies of credit transferability have shown that credit loss is widespread and has negative effects on degree attainment (e.g., Monaghan & Attewell, 2015). Comparing credit loss with other barriers community college students encounter in seeking a bachelor's degree, Monaghan and Attewell found that loss of credits upon transfer was a substantial barrier to degree completion. Students who were able to transfer most of their credits were 2.5 times more likely to complete a bachelor's degree than those who were not.

Studies of credit transferability likely underestimate the inefficiency of the credit transfer process since the *applicability of transfer credit* is often ignored. For students to transfer their credits efficiently, not only does the receiving institution have to accept their credits for general or "elective" credit, but those credits must also apply toward students' major requirements. While applicability of transfer credit is a more accurate measure of credit transfer efficiency than is credit transferability, researchers have rarely used it to evaluate students' transfer outcomes, in part due to the complexity of mapping student transcripts to degree requirements, which frequently vary by institution and even by department and can change from year to year. Ideally, credit applicability should be measured through audits of transcripts of students who transferred and earned a bachelor's degree from a given university. Because it is difficult for outside researchers to gain access to such data from multiple institutions, such studies are rare. One exception is a 2001 report by the Transfer Issues Advisory Committee convened by the Texas Higher Education Coordinating Board, which was based on transcript audits conducted by five Texas universities. That report, which looked at students who had earned a bachelor's degree after transferring from a community college, revealed that 83 percent of credit hours presented by transfer students who had earned at least 30 credits hours at a

Texas community college were accepted for transfer, but only 70 percent of the credits were accepted as applicable toward a bachelor's degree in the students' majors.[1]

Another measure researchers have used to study credit transfer efficiency is *excess credits among completers*, calculated as the total number of credits earned or attempted by graduates beyond those required for a particular degree. There are numerous examples of statewide analyses of excess credits among degree completers. Complete College America (2011, n.d.) detailed average credits accumulated by bachelor's degree completers in more than 30 states, reporting that on average students completed a bachelor's degree with 135 credits. Other research suggests that bachelor's degree completers who start at a community college earn even more excess credits on average than do those who start at a four-year institution. Using data from the Texas Higher Education Coordinating Board, Cullinane (2014) found that community college transfers who earned a bachelor's degree attempted 150 college credits, whereas students who started at a four-year institution attempted 142 credits. Cullinane's results did not include any remedial credits students attempted, even though the majority of students who enter community colleges take at least one remedial course—so the overall number of excess credits attempted was likely even higher. Cullinane also matched equivalent groups of community college transfer students and native four-year students and found that transfers were less likely to complete a bachelor's degree; those who did took longer to complete. In a similar study, Xu, Jaggars, and Fletcher (2016) examined matched samples of new college students who started at a two- or four-year college in Virginia and indicated an intent to earn a bachelor's degree. Xu et al. found that on average two-year entrants who completed a bachelor's degree earned 10 more credits than did similar four-year entrants, and they took two semesters longer to graduate.

[1] Our analysis of that report indicates that, of the 30 percent of transfer credits not applied toward a bachelor's degree program, over 75 percent were rejected for reasons that seem unclear: other reasons (51 percent), designated "technical" courses (12 percent), or no course level equivalents (12 percent). Less than 25 percent were rejected because of low grades (9 percent), because they were from developmental courses (7 percent), or for other clear reasons. This suggests that there were no clear reasons why one in every five community college credits students tried to transfer were not accepted for credit toward a degree in their major.

1.2 Understanding and Addressing Credit Transfer Inefficiency

For many community college students who desire to transfer to a four-year institution and earn a bachelor's degree, there is no clear path. In an analysis of National Student Clearinghouse data on around 100,000 students who started at a community college in fall 2007, transferred to a four-year institution, and earned a bachelor's degree within six years, only 8 percent followed a "2 + 2" pattern of spending two years at a community college followed by two years at a four-year institution (Fink, 2017). Another study using National Student Clearinghouse data found that, taking into account term-by-term enrollment at a community college or a four-year college, among the 3,290 students who enrolled at a community college in 2007 and earned a bachelor's degree in computer science within seven years, there were 1,213 unique term-by-term patterns (Jaggars, Fink, Fletcher, & Dundar, 2016).

The remarkable variation among states and individual community colleges in the average rate at which students earn community college degrees before transferring, as reported in Jenkins and Fink (2016), is another indicator of the lack of clear transfer pathways. In a follow-up study, Wyner, Deane, Jenkins, and Fink (2016) exploited this variation in individual colleges' transfer outcomes to identify and study pairs of two- and four-year colleges that had higher-than-expected bachelor's completion rates for students who started at a community college and later transferred, controlling for student and institutional characteristics.[2] In looking at the practices common to these high-performing transfer partnerships, Wyner et al. emphasized the role of program-level curricular alignment between community colleges and universities in helping students take the right sequence of courses to maximize credit transfer applicability toward their desired majors and to minimize overall excess credits. Successful partner colleges had detailed transfer guides showing students which courses to take to ensure that all of their credits count toward their desired major at the four-year destination college. The university partners encouraged community colleges to help students explore and select a major or broad field of interest soon after entry so they could take any lower division coursework required for that major or field of interest and avoid having to take other prerequisites later.

[2] See also Fink and Jenkins (2017).

To increase transfer student success and address credit transfer inefficiency, state and college leaders are attempting to map out curricular paths for transfer students more clearly. Baker (2016) evaluated California's Associate Degrees for Transfer (ADTs), which are statewide, major-specific agreements between the California Community Colleges and California State Universities (CSUs). The ADTs were designed to increase transfer to the CSUs by giving students a structured curriculum at the California Community Colleges that would prepare them for more efficient transfer to the CSUs. Baker took advantage of the phased rollout of these structured agreements across colleges and departments to assess (using a "differences-in-differences-in-differences" approach) the effects of the ADTs on degree completion and transfer. While Baker did not look at credit transfer efficiency directly, Baker found that the ADTs resulted in more associate degree completions and had a marginal effect on the rate of transfer to the CSUs.

Washington State has also introduced statewide, structured transfer pathways to increase rates of transfer and credit transfer efficiency. In 2011, the Washington State Higher Education Coordinating Board studied excess credits earned by community college transfer students who used different statewide transfer agreements along the path toward a bachelor's degree in business (Washington State Higher Education Coordinating Board, 2011). While the study was descriptive and did not account for student background characteristics, the results showed that students who followed the business major–specific transfer agreement and graduated earned eight fewer excess quarter credits than did students who followed a generic, field-independent transfer agreement that gave them flexibility to choose among many "general education" courses from a distribution list. Furthermore, students following the major-specific transfer agreement graduated with 11 fewer excess quarter credits than did students who did not follow any transfer agreement. These findings underscore the importance of mapping out curricular paths from community college through completion at a four-year institution in order to decrease credit transfer inefficiency.

1.3 Using Data Mining to Analyze the Sources of Inefficiency in Complex Curricular Pathways

Building on the work of Clifford Adelman (2005, 2006), many higher education researchers have sought to analyze student transcript data to understand student characteristics and behaviors and institutional practices associated with efficient degree completion. Community college researchers have used this approach extensively (Bahr, 2013; Belfield, Crosta, & Jenkins, 2016; Bragg, 2012; Calcagno et al., 2007; Hagedorn, 2005; Leinbach & Jenkins, 2008). Bahr (2013) further described this approach as deconstructing students' academic pathways through in-depth analyses of their coursework in order to surface structural barriers students encounter while pursuing their goals. Using transcripts to deconstruct and analyze student academic pathways is essential for uncovering inefficiencies in community college students' academic pathways to a baccalaureate (Hagedorn & Kress, 2008).

Most of the research deconstructing student transcripts has relied on regression analysis and other conventional statistical methods to test hypotheses about the factors associated with positive student outcomes. Recent advances in data mining address limitations of using linear statistical analysis to understand complex social science datasets, such as overlooking non-linear relationships and complex interaction effects (Attewell & Monaghan, 2015). Given the complexity of course-taking patterns observed in student transcript databases, data-mining techniques are particularly well suited for the analysis of associations between course-taking patterns and student outcomes such as the efficiency of credit transfer. In a recent study, Wang (2016) demonstrated how data-mining techniques could be used with student transcript data to describe the course-taking behaviors of community college students who transferred in science, technology, engineering, and mathematics (STEM) fields versus non-STEM fields. Wang found that among STEM transfer students, contrary to conventional wisdom, the early course-taking patterns that were most associated with upward transfer in STEM involved taking transferable STEM courses during the first term followed by math sequences in subsequent terms. This pattern suggests that students may benefit from getting exposure to a discipline before they decide to tackle challenging math courses and other requirements for the given field.

In this study, we use data mining and a rich set of student transcript data to better understand course-taking patterns and other behaviors among community college students who transferred and earned a bachelor's degree with more or fewer excess credits. Specifically, we measure the number of credits attempted beyond the minimum required for bachelor's degree programs by students in two states who started at a community college, transferred, and earned a bachelor's degree. We then examine the course-taking patterns of community college entrants who transferred to a four-year college and obtained a bachelor's degree and the course-taking patterns of a sample of native four-year entrants who completed a bachelor's degree to identify what enrollment patterns and other factors may be associated with the success of transfer students who have accumulated fewer excess credits. Finally, we examine how well the course-taking indicators identified through the data-mining analysis explain the variance in students' excess credits.

2. Data

The data for this study are taken from transcripts of first-time-in-college students in two states, referred to as State A and State B, who earned a bachelor's degree from a public university in the given state within six years (see Table 1). Students from State A began at one of the state's public two- or four-year institutions between the fall of 2008 and the spring of 2009, and students from State B started at one of the state's community colleges between the fall of 2004 and the summer of 2006.[3] Importantly, our sampling strategy requires all community college entrants to transfer to a public four-year institution within the time period observed in order to be included in the analysis. Excluded from all samples are the small number of students for whom we cannot account for more than 30 percent of the credits required for graduation (or students with fewer than 80 college-level credits).[4] The final analytic samples include 666 community college entrants and 5,158 four-year college entrants from State A, and 12,722 community college entrants from State B.

[3] There are two cohorts of community college entrants in State B: The 2004 cohort started between the fall of 2004 and the summer of 2005, and the 2005 cohort started between the fall of 2005 and the summer of 2006. All students in State B were tracked until the spring of 2010.
[4] Students may have earned credits outside of the given state system. Less than 1 percent of students were excluded from the final analysis based on this restriction. Other cases where student credits were not accounted for in the state dataset resulted in 5–10 percent of students in the sample graduating with fewer than 120 credits attempted. We recoded these students' excess credits from negative values to zero.

Table 1
Six-Year Transfer and Graduation Outcomes

College System	*n*	Transferred to an In-State Public Four-Year College	Earned a Bachelor's Degree	Earned a Bachelor's Degree With 80+ Credits
State A, four-year	10,844		47.8%	47.6%
State A, two-year	9,944	14.0%	6.7%	6.7%
State B, two-year	174,749	18.8%	7.5%	7.3%

The datasets provided by the state systems include a rich set of demographic information, including sex, age, and race/ethnicity as well as proxies for academic ability as determined by high school grade point average (GPA) and enrollment in developmental education courses. Table 2 reports demographics and academic characteristics for each subsample. Compared with four-year entrants, students who began at a community college in State A are more likely to be White, female, and older; they are also more than three times as likely to take developmental education courses. Generally, community college entrants in State B are demographically similar to those in in State A, though they are more likely to be racial/ethnic minorities. State A and State B also differed in the number of credits attempted and completed by community college entrants pre-transfer. On average, students who began at a community college in State A attempted 70 semester credits and earned 60 credits before they transferred to a four-year college in the state, whereas students in State B attempted 39 semester credits and earned 35 credits pre-transfer. This indicates that community college entrants in State B generally transferred earlier in their academic careers than did those in State A.

Table 2
Sample Characteristics

Characteristic	State A, Four-Year Entrants	State A, Two-Year Entrants	State B, Two-Year Entrants
White	71%	83%	75%
Female	58%	62%	57%
Average age	18.6	19.6	19.3
Pell recipient	17%	17%	20%
Attempted developmental education	17%	58%	61%
Average credits attempted pre-transfer (*SD*)		69.7 (22.4)	39.1 (32.5)
Average credits earned pre-transfer (*SD*)		60.1 (20.4)	35.3 (31.4)

2.1 Outcome Variables

We conducted our analyses using the total number of excess credits attempted and earned as outcome variables, but given the implications for efficiency and cost—as students and taxpayers pay for all courses students attempt whether or not they complete them—we report results using the number of excess credits attempted (including both developmental and college-level credits[5]). We calculated the number of excess credits by subtracting the number of credits attempted from the number of credits required to complete the bachelor's degree program from which a student graduated. Information on required credits by program was obtained from each university's academic catalog. Analyses using the number of excess college-level credits earned as the outcome variable yielded similar results; therefore, we focus our reporting here on the analysis of credits attempted for parsimony.[6]

As shown in Table 3, the number of total excess credits attempted among community college entrants who transferred and earned a bachelor's degree from a state university is similar in State A and State B: 27 and 29 credits, respectively. On average, bachelor's degree completers who started at a public four-year institution in State A attempted fewer excess credits (M = 19.5, SD = 18.0) than did bachelor's degree completers who started at a community college (M = 28.8, SD = 20.3). Further examination of excess credits attempted across the three subsamples reveals differences by student characteristics. White students, Hispanic students, and students whose race/ethnicity is unknown tended to be at or below the average number of excess credits attempted, whereas Black students, American Indian students, and Asian students tended to attempt above-average numbers of excess credits (see Table 3). In State A, Black students who started at a community college, transferred, and earned a bachelor's degree attempted on average 46 credits above the number required for their degree. In State B, the average number of excess credits attempted by Black transfer students who earned a bachelor's degree was 36. Additionally, male students and Pell grant recipients attempted more excess credits on average than did female students and non-Pell recipients across the three subsamples, and students with more developmental education placements attempted more excess credits.

[5] We performed a robustness check using only college-level credits attempted and found similar, though attenuated, results.
[6] Results from analyses of excess credits earned are included in Appendix B.

Among community college entrants who went on to transfer and earn a bachelor's degree in State A, there was only about a two-credit difference in the average number of excess credits attempted by associate degree completers and by students who did not complete an associate degree. In State B, however, community college entrants who completed an associate degree prior to completing a bachelor's degree attempted about seven more excess credits on average than students who did not complete an associate degree.

Table 3
Average Excess Credits Attempted Among Bachelor's Degree Completers

	State A, Four-Year Entrants (*n* = 5,158)		State A, Two-Year Entrants (*n* = 666)		State B, Two-Year Entrants (*n* = 12,722)	
Characteristic	*M*	*SD*	*M*	*SD*	*M*	*SD*
All students	19.5	18.0	28.8	20.3	27.3	22.9
Race/ethnicity						
White	17.3	16.8	28.0	20.0	25.3	21.8
Unknown	22.3	18.5	27.3	18.1	27.4	22.7
Hispanic	19.6	18.1	22.8	17.2	28.3	23.2
Black	26.1	20.0	45.8	23.0	36.3	25.6
American Indian	26.2	13.4	35.0	7.1	35.4	28.1
Asian	23.7	17.4	34.6	20.0	30.9	24.7
Gender						
Female	18.5	17.4	26.8	19.6	26.0	22.8
Male	20.9	18.7	32.3	21.1	29.1	23.0
Age						
< 18	20.5	17.0	23.2	15.0	26.2	21.0
18–24	19.4	18.0	29.1	20.6	29.2	24.2
25–34	25.1	18.5	26.3	18.5	21.7	24.1
35–44	22.9	19.9	26.8	17.1	23.2	23.8
45+	27.3	1.5	29.6	11.1	26.8	29.0
Pell recipient						
Yes	24.5	19.7	29.7	21.4	34.8	25.4
No	18.5	17.4	28.6	20.1	25.9	22.2
Developmental education placements						
Zero	19.5	17.9	24.5	21.5	22.3	20.3
One	44.0	16.6	27.6	18.8	22.0	20.8
Two			35.3	18.3	30.7	23.4
Three			40.8	15.8	47.1	24.7
Associate degree completed						
Yes	35.7	22.1	28.1	18.6	32.9	25.0
No	19.4	17.9	29.7	22.0	25.9	22.2

2.2 Course-Taking Variables

We used longitudinal transcript data to examine numerous variables representing course-taking behaviors of interest over time. Specifically, we calculated the proportion of total credits attempted and earned at several time points. Course-taking indicators for four-year entrants were calculated before and after students accumulated 60 college-level credits, which is an important threshold, as it is generally considered the point at which students should have entered major-specific coursework. For community college entrants, in addition to looking at course-taking before and after the 60-credit threshold, we computed course-taking behaviors prior to transfer to a four-year institution, during the first term following transfer, and during the first two non-summer terms following transfer. Table 4 lists the thresholds at which course-taking indicators were calculated for two- and four-year entrants.

Course-taking behaviors of interest include the proportion of credits attempted and earned within each time frame of interest that were:

- 100-, 200-, or 300-level (or higher);
- in the same academic area as the student's bachelor's degree; and
- designated as STEM courses (data only available in State A).[7]

Table 4
Thresholds of Credit Accumulation

Timing	Two-Year	Four-Year
Pre-transfer	x	
Post-transfer	x	
First term post-transfer	x	
First two non-summer terms post-transfer	x	
Prior to accumulating 60 credits	x	x
After earning at least 60 credits	x	x

[7] See https://www.ice.gov/sites/default/files/documents/Document/2016/stem-list.pdf.

In addition, we included indicators for the percentage of total and college-level credits attempted and earned at a community college, as courses taken at the community college can vary in transferability and applicability. Although similar information was not available for State A, for State B we were able to calculate the proportion of credits earned at a community college prior to transfer that were part of the statewide transfer course library.

Looking at proportions of credits earned or attempted provides a more standardized measure of course-taking behavior than the count of credits earned or attempted by effectively controlling for full-time or part-time status. Calculating proportions of credits earned or attempted also accounts for potential differences in the number of credits awarded for a given course. Importantly, we considered whether or not a student attempted a 100-level course in English and math during each period of interest because these courses often serve as program "gatekeepers" and can indicate that a student has yet to fulfill lower division prerequisites. For community college entrants we also included a series of indicators for whether the student completed different components (subject areas) of the general education curriculum and whether the student earned an associate degree or certificate prior to transfer. The final dataset, comprising all demographic covariates and measures of course-taking behavior, included 107 variables for analysis in State A and 67 variables for analysis in State B. The full list of variables is included in Appendix C.

2.3 Analytic Plan

Given that the purpose of the study is to explore differences in characteristics and course-taking patterns among students with more or fewer excess credits, we used partition trees, a data-mining technique, to examine the relationship between course-taking behaviors and other factors and credits attempted by students who earned bachelor's degrees. Using partition trees offers two key advantages for this study compared with other techniques such as regression analysis (Attewell & Monaghan, 2015, pp. 162–163). First, the ratio of independent variables to observations is less important in partition tree analysis than in regression analysis, where models are unreliable with fewer than 15 to 30 observations per variable. In the current study, we identified more than 100 variables to capture students' course-taking patterns for use with samples as small as 600 observations.

The second, and perhaps more important, advantage of partition trees is that they do not simply estimate average relationships between the independent and dependent variables. Rather, partition trees recursively classify observations based on the variable that best partitions the data at each stage. The result is a complex classification scheme that can reveal unforeseen interactions between independent variables. This type of result is ideal output for an exploratory study of this sort.

While partition trees are powerful tools for exploratory analysis, the results from partitioning the data can be difficult to interpret. A partition tree may automatically partition a particular dataset many times before it can no longer better partition any remaining group of observations. Although the result of this partitioning might explain variance in a given outcome very accurately, it can be difficult to interpret such a detailed classification scheme. Simpler schemes, on the other hand, may be less accurate, so choosing a method of analysis involves a trade-off between accuracy and interpretability (Attewell & Monaghan, 2015).

In our analyses, we used output from partition tree analyses to understand the overall ability of our independent variables to predict the number of students' excess credits and to examine the course-taking patterns among students who accumulated more or fewer excess credits. First, we partitioned the data until the tree was optimized, such that additional partitions would not improve predictive power as measured by *R*-squared (R^2). In this analysis we also considered the most important variables in partitioning the data (or the "column contributions" from the model) in the full model. While this ranking of predictor variables does not specify if a variable has a positive or negative relationship with the outcome, it does show the relative importance of that variable compared with others included in the analysis. The results from this procedure allow us to understand how well each variable predicts the outcome and which variables were most important in partitioning the data.

After examining the model summary and column contributions for a full partition tree, we visually examined a simplified version of the tree that shows only the initial partitions. This step allows for a more nuanced understanding of how students were first classified into lower and higher excess credit groups. The classifications in the partition tree indicate the directionality of a given predicting variable, though beyond the first partition these classifications are only germane to specific subgroups of students. Thus, information from different groups of students (e.g., four-year entrants, two-year entrants, students from different states) and different analyses (e.g., model summaries, column contributions, visual examination of simplified partition trees) must be triangulated in order to understand the course-taking patterns among bachelor's completers with more and fewer excess credits.

Finally, we used regression models to examine how well the variables of interest identified through our exploratory data-mining analyses explain students' excess credits. The use of regression models complements the data-mining analysis in that it also allows us to compare varying sets of independent variables in terms of how well each explains variance in the outcome variable.

3. Results

3.1 Partition Tree Analysis

Four-year entrants, State A. As shown in Table 5, the partition tree analysis of excess credits attempted among four-year entrants from State A resulted in a total of 16 splits with R^2 values of .295 and .233 for the training and validation samples, respectively. The table also shows the column contributions, with the number of times each variable was used to split (partition) the data. For each of the column contributions Table 5 shows the portion of the sum of squared errors (*SSE*) each variable accounted for in the model, which we used to rank the variables' importance in the classification scheme. Finally, based on visual examination of the simplified partition trees, we have included a column in Table 5 to aid in the interpretation of our results by indicating if higher values for the predictor generally resulted in more or fewer excess credits.

Variables describing the percentage of 300-level credits earned before or after the student had accumulated 60 college-level credits represented two of the three most important predictors of excess credits attempted among four-year entrants in State A. The other top predictor was an indicator for whether or not a student attempted a 100-level math course after the 60-credit threshold. Looking at the simplified tree (Figure A1), students were first split by whether, after 60 credits, they earned more or less than 16.7 percent of their credits at the 300 level or higher. After reaching the 60-credit threshold, students who earned a higher proportion of 300-level credits attempted fewer excess credits (*M* = 11.8, *SD* = 14.0) than did students who earned less than 16.7 percent of their credits at the 300 level (*M* = 24.8, *SD* = 18.3). Students who earned 30.4 percent or more of their credits in 300-level courses after 60 credits attempted even fewer excess credits (*M* = 7.8, *SD* = 11.8). Among students who earned less than 16.7 percent of their credits in 300-level courses after 60 credits, those who also attempted more than 11.2 percent of their credits in 200-level courses during that time attempted more excess credits overall (*M* = 28.9, *SD* = 18.9), and those who also took a 100-level math course during that time attempted even more excess credits (*M* = 37.8, *SD* = 19.9).

Table 5
Model Summary and Column Contributions:
State A, Four-Year Entrants, Excess Credits Attempted

	R^2	*n*	Result of First Split on Excess Credits[1]	Number of Splits	Portion of *SSE*
Training sample	.295	3,867		16	
Validation sample	.233	1,291			
Variable					
Percentage of credits earned at 300+ level after 60 credits			Fewer excess	2	0.4156
Attempted 100-level math after 60 credits			More excess	1	0.1696
Percentage of credits earned at 300+ level before 60 credits			More excess	3	0.1485
Percentage of credits attempted at 300+ level after 60 credits			Fewer excess	1	0.0791
Percentage of credits attempted at 300+ level before 60 credits			More excess	2	0.0496
Percentage of credits attempted at 200 level after 60 credits			More excess	1	0.0346
Percentage of credits earned in STEM courses before 60 credits			Fewer excess	1	0.0274
Race/ethnicity: White			++	1	0.0244
Race/ethnicity: Black			++	1	0.0213

[1] This column describes the directionality of the first time a given variable was used to partition the data. Higher values (e.g., larger percentages, higher GPAs) resulted in either more or fewer excess credits on average, as indicated in this column.

++ Variables used to partition data beyond the eighth split are not visually examined.

Two-year entrants, State A. As shown in Table 6, the partition tree of excess credits attempted among two-year entrants from State A resulted in a total of 11 splits, with R^2 values of .511 and .195 for the training and validation samples, respectively. Unlike other partition tree models in this study, the validation sample R^2 was much lower than the R^2 of the training sample, which might be explained by the smaller number of students in this subsample. Similar to results for the four-year entrants, two of the top three predictors were variables describing the percentage of 300-level courses students took before or after the 60-credit threshold. Additionally, the number of developmental education placement areas and the proportion of 100-level courses students took in the first two terms after transfer were identified as important in partitioning the data. Examining the simplified partition tree (Figure A2), the first split grouped students by whether they attempted more or less than 7.3 percent of their credits in 200-level courses after 60 credits: Students who attempted fewer credits in 200-level courses averaged 16.2 excess credits attempted (SD = 11.9) compared with other students (M = 34.6, SD = 21.2). Students who attempted fewer credits in 200-level courses after 60 credits and earned less than 55 percent of their credits at a community college averaged 12.2 excess credits attempted (SD = 11.0), and those who also took less than 3 percent of their credits in STEM courses after 60 credits had even fewer excess credits attempted (M = 5.5, SD = 5.1). Students who, after 60 credits, took more than 7.3 percent of their credits in 200-level courses and attempted more than 10 percent of their credits in STEM courses averaged 41.1 excess credits attempted (SD = 21.0). Those who also attempted less than 14.3 percent of their credits in 300-level courses after passing the same threshold had even more excess credits attempted (M = 42.7, SD = 20.3). In short, students who after 60 credits took more 200-level credits, more STEM credits, and fewer 300-level credits tended to have more excess credits attempted.

Table 6
Model Summary and Column Contributions: State A, Two-Year Entrants, Excess Credits Attempted

	R^2	n	Result of First Split on Excess Credits[1]	Number of Splits	Portion of *SSE*
Training sample	.511	510		11	
Validation sample	.195	156			
Variable					
Percentage of credits attempted at 300+ level after 60 credits			Fewer excess	2	0.5572
Number of developmental education placement areas			More excess	3	0.1572
Percentage of credits attempted at 300+ level before 60 credits			++	1	0.0893
Percentage of credits attempted at 100 level in two terms after transfer			More excess	1	0.0774
Percentage of credits attempted in STEM after 60 credits			More excess	1	0.0437
Percentage of credits earned in STEM after 60 credits			++	1	0.0347
Percentage of credits earned at 300+ level in two terms after transfer			++	1	0.0212
Percentage of credits attempted at 200 level after 60 credits			More excess	1	0.0193

[1] This column describes the directionality of the first time a given variable was used to partition the data. Higher values (e.g., larger percentages, higher GPAs) had either more or fewer excess credits on average, as indicated in this column.

++ Variables used to partition data beyond the eighth split are not visually examined.

Two-year entrants, State B. As shown in Table 7, the partition tree of excess credits attempted among two-year entrants from State B resulted in a total of 135 splits, with R^2 values of .581 and .510 for the training and validation samples, respectively. While it is challenging to interpret many of the variables identified as important in partitioning the data given the large number of splits, we observed the directionality of the first split for four of the top six variables. Students who took more courses in the statewide transfer library attempted fewer excess credits. Students who earned more credits in 100-level courses before 60 credits, attempted a 100-level course after 60 credits, and attempted more 200-level courses after 60 credits attempted more excess credits. As in State A, 100-level courses taken after 60 credits and the number of developmental education placements were identified as important in partitioning the data. Examining the simplified partition tree (Figure A3), the first split separated a group of students with fewer excess credits attempted overall (*M* = 15.3, *SD* = 17.2) who attempted more than 2 percent of their credits in 300-level courses before 60 credits from those who attempted less than 2 percent (*M* = 33.5, *SD* = 23.3). Among the former, those who also attempted more than 2 percent of their credits in 300-level courses one year after transfer had fewer excess credits attempted (*M* = 10.6, *SD* = 14.4) than those who did not (*M* = 20.9, *SD* = 18.5); among the latter, those who earned less than 86 percent of their credits in transferable courses averaged 38.9 excess credits attempted (*SD* = 24.0), versus 26.2 for other students (*SD* = 20.3).

Table 7
Model Summary and Column Contributions:
State B, Two-Year Entrants, Excess Credits Attempted

	R^2	*n*	Result of First Split on Excess Credits[1]	Number of Splits	Portion of *SSE*
Training sample	.581	9,516		135	
Validation sample	.510	3,206			
Variable					
Percentage of credits attempted in 200-level courses at a community college			++	32	0.1979
Percentage of credits earned in 100-level courses before 60 credits			More excess	1	0.1935
Percentage of credits earned in transferable courses			Fewer excess	6	0.1539
Percentage of credits attempted at a community college			++	9	0.0750
Percentage of credits attempted in 200-level courses after 60 credits			More excess	2	0.0745
Took 100-level math after 60 credits			More excess	1	0.0518
Number of developmental education placement areas			++	5	0.0468
Percentage of credits attempted in 100-level courses at a community college			++	4	0.0417
Took 100-level math before 60 credits			++	1	0.0304
Percentage of credits attempted in 100-level courses after 60 credits			++	2	0.0204
Percentage of credits attempted in transferable courses			++	3	0.0176
Race/ethnicity: Black			++	2	0.0156
Took 100-level English before 60 credits			++	4	0.0156
Percentage of credits earned in 200-level courses at a community college			++	3	0.0143
Percentage of credits attempted in 100-level courses one year after transfer			++	2	0.0127
Percentage of credits attempted in 300-level courses one year after transfer			Fewer excess	1	0.0089
Completed general education requirement at a community college			++	2	0.0089
Percentage of credits attempted in degree area at a community college			++	1	0.0077
Race/ethnicity: White			++	1	0.0043
Percentage of credits attempted in 100-level courses one term after transfer			Fewer excess	1	0.0029
Percentage of credits attempted in 200-level courses one year after transfer			++	1	0.0029
Two-digit CIP code for major			++	1	0.0011
Age at first enrollment			++	1	0.0010
Percentage of credits attempted in 200-level courses before 60 credits			++	1	0.0006

[1] This column describes the directionality of the first time a given variable was used to partition the data. Higher values (e.g., larger percentages, higher GPAs) had either more or fewer excess credits on average, as indicated in this column.

[++] Variables used to partition data beyond the eighth split are not visually examined.

3.2 Partition Tree Results Summary

The three partition tree models across the two states varied in their level of complexity and predictive power. The three partition trees ranged in complexity from requiring 11 to 135 splits before reaching optimal predictive power. The predictive power of the training samples, as measured by R^2 values, ranged from .30 to .58. The predictive power of the validation samples nearly matched that of the training samples with the exception of the model for two-year entrants in State A, which may be explained by the smaller sample size. The full partition trees utilized complex interactions among variables to attain their predictive power.

Despite this variation, the main findings were strikingly similar between the models for two- and four-year institutions in State A and between the models for two-year institutions in State A and State B. To more easily interpret the tree results, we examined simplified trees that only included the first levels of partitioning (see Appendix A). Bachelor's completing four-year entrants who, after accumulating at least 60 college-level credits, took more 300-level courses, took fewer 200-level courses, and did not have to take a 100-level math course attempted fewer excess credits. Bachelor's completing two-year entrants in State A attempted fewer excess credits if they took fewer 200-level courses after 60 credits and if they took less than 55 percent of their credits at a community college. These students also attempted fewer excess credits if, after 60 credits, they took a smaller percentage of STEM courses and a larger percentage of 300-level courses. Students attempted even fewer excess credits if, in addition to having the right balance of 200- and 300-level courses after the 60-credit threshold, they accumulated less than 63 percent of their credits in 100-level courses before hitting the 60-credit threshold. In State B, bachelor's completing two-year entrants attempted fewer excess credits if they took more 300-level courses both after 60 credits and within one year after transfer and, encouragingly, if they took most of their courses from the state transfer library.

The simplified partition trees also illustrate subgroups of students who attempted many more excess credits on the way to earning a bachelor's degree than the overall average. For example, the 666 two-year entrants who completed a bachelor's degree in State A attempted 29 excess credits on average. However, a subgrouping of 155 of those students averaged 46 excess credits attempted, and these students shared the following course-taking behaviors: After 60 credits, they earned more than 9 percent of their credits in 200-level courses, less than 14 percent of their credits in 300-level courses, and more than 10 percent of their credits in STEM courses. In another example, among the 12,722 two-year entrants who completed a bachelor's degree in State B with an average of 27 excess credits attempted, a subgrouping of 3,648 students who took less than 86 percent of their courses from the statewide transfer library and attempted less than 2 percent of their first 60 credits in 300-level courses averaged 39 excess credits attempted.

3.3 Examining Explanatory Effects of Partition Tree Findings

Analytic plan. To complement the descriptive findings from the partition tree analyses, we conducted supplemental regression analyses for each sample to understand the relative explanatory effects of each set of variables identified as being associated with excess credits in the data-mining analyses. We grouped the independent variables in the regression analyses into three sets. In the first set we included student characteristics, such as demographics, high school GPA, and developmental education placement, in order to control for these variables as we added in the second and third sets. In the second set we included a relatively simple group of course-taking variables that, taken together, measure the percentage of student coursework taken at the 100 or 300 levels. These simplified transcript variables were included because they were identified as important predictors of excess credits globally across our samples in the partition tree analysis. Additionally, these variables capture whether students progressed from introductory (100-level) to advanced (300-level) courses as they accumulated more credits. In the third set we included variables identified from each sample's partition tree analysis. Based on the first six splits in the partition tree, we selected six variables and used the cut points identified in the partition tree analysis to create six dummy variables.[8] For each sample, we ran three regression models to examine how the results changed with the addition of each set of variables. Furthermore, we ran a fourth model for each sample to compare the explanatory power between variables included in the second set (simpler transcript

[8] Due to sample size restrictions and the total number of variables included in the regression models, we limited the number of variables in the third step to the top six splits. Despite the sample size being larger in State B, we maintained the same number of variables across the samples for comparability.

variables) and the third set (variables identified through data mining). Given the exploratory nature of this study, beta coefficients significant to the level of $p < .01$ were considered meaningful in the regression models.

Results summary. Tables 8–10 present results from the supplemental regression analyses for each of the samples, and the findings from these analyses are summarized in Table 11. The full model (Model 4) explained 30 percent of the variance in excess credits attempted for four-year entrants in State A ($F = 104.42$, $p < .001$), 46 percent of the variance in excess credits attempted for community college entrants in State A ($F = 26.44$, $p < .001$), and 30 percent of the variance in excess credits attempted for community college entrants in State B ($F = 249.44$, $p < .001$). Models 2 and 3, which included indicators from student transcripts, reliably improved the explanatory power of the regression models across all samples compared with the explanatory power of the set of student characteristic variables included in Model 1. Across the three samples there was similarity in the explanatory power of Model 2, which included student characteristics and simple transcript indicators, and that of Model 3, which included student characteristics and the data-mined cut-point transcript indicators. In other words, results across the three samples suggest that the simple transcript indicators have about the same ability to explain the variance in excess credits attempted as the indicators identified through the partition tree analysis, controlling for student characteristics.

Table 11 also summarizes the independent variables significantly associated with excess credits attempted in the full model (Model 4) across the three samples. The beta coefficients in Tables 8–11 indicate the additional amount of excess credits attempted that are associated with each student characteristic or transcript indicator. Notably, being Black was associated with attempting 6–8 more excess credits than being White in the State A four-year and State B two-year samples. Additionally, among community college entrants in State B, being Asian or American Indian was associated with attempting 4–6 more excess credits, completing an associate degree was associated with attempting five more excess credits, and each developmental education placement was associated with attempting another six excess credits. Across the three samples, taking more 300-level courses before reaching the 60-credit threshold was associated with attempting fewer excess credits. For community college entrants in State B, taking more 100-level courses before and after the 60-credit threshold was associated with attempting more excess credits. Although earning a higher percentage of credits at the community college was negatively associated with excess credits attempted among community college entrants in both states, four-year entrants in State A who earned more of their credits from a community college attempted more excess credits overall. It may be that students transferring credits inefficiently are required to take extra courses at the four-year college, adding excess credits and further reducing the proportion of community college credits among all credits earned.

The data-mined cut-point indicators give a more precise estimate of the effects of particular course-taking behaviors on excess credits (see Table 11 for a summary). Some of the largest positive associations with excess credits attempted were taking a 100-level math course after 60 credits (8.0 more excess credits, State A four-year sample), taking more than 2 percent of coursework prior to transfer in 200-level courses (6.65 more excess credits, State B two-year sample), taking more than 10 percent of coursework after 60 credits in 200-level courses (7.42 more excess credits, State B two-year sample), and taking more than 44 percent of coursework before 60 credits in 100-level courses (2.46 more excess credits, State B two-year sample). Negative associations with excess credits attempted included taking more than 87 percent of coursework from the statewide transfer library (7.29 fewer excess credits, State B two-year sample), taking more than 71 percent of coursework after 60 credits in 300-level courses (8.04 fewer excess credits, State A two-year sample), and completing more than 45 percent of overall coursework at a community college (9.65 fewer excess credits, State A two-year sample).

Table 8
Multiple Regression Analysis: State A, Four-Year Entrants, Excess Credits Attempted

Variable	Model 1	Model 2	Model 3	Model 4
Student characteristics				
Pell recipient	3.593**	2.280**	2.018**	1.816**
	(0.673)	(0.599)	(0.596)	(0.585)
Race/ethnicity				
Unknown	4.129**	2.854**	2.453*	2.463*
	(1.234)	(1.098)	(1.090)	(1.071)
Hispanic	1.733	2.033	1.723	2.037
	(1.637)	(1.456)	(1.445)	(1.420)
Black	8.099**	6.825**	5.764**	5.969**
	(0.640)	(0.571)	(0.574)	(0.564)
American Indian	8.794*	5.366	4.086	4.913
	(5.263)	(4.675)	(4.645)	(4.560)
Asian	6.032**	3.374*	3.529**	3.231*
	(1.539)	(1.370)	(1.360)	(1.336)
Female	-3.026**	-1.664**	-2.224**	-1.866**
	(0.494)	(0.441)	(0.438)	(0.431)
Age	0.248	0.252*	0.163	0.175
	(0.139)	(0.124)	(0.123)	(0.121)
Number of developmental education placements	13.701**	10.409*	12.155**	9.353*
	(4.528)	(4.100)	(3.996)	(4.000)
Completed associate degree	16.313**	0.097	8.858**	-0.018
	(2.659)	(2.648)	(2.358)	(2.586)
Simple transcript indicators				
Percentage of all credits attempted at a community college		34.269**		31.276**
		(6.457)		(6.308)
Percentage of credits before 60-credit threshold attempted at 100 level		3.666		4.577*
		(2.094)		(2.097)
Percentage of credits after 60-credit threshold attempted at 100 level		0.131		-8.429*
		(3.321)		(3.313)
Percentage of credits before 60-credit threshold attempted at 300 level		-35.126**		-23.203**
		(2.322)		(3.943)
Percentage of credits after 60-credit threshold attempted at 300 level		-42.907**		-36.098**
		(2.218)		(3.459)
Data-mined cut points and variables				
Percentage of credits after 60-credit threshold attempted at 300 level: > 75%			-6.795**	-0.731
			(0.516)	(0.751)
Percentage of credits before 60-credit threshold attempted at 300 level: > 14%			-3.945**	-1.152
			(0.674)	(0.758)
Percentage of credits before 60-credit threshold attempted at 300 level: > 18%			-4.343**	-1.553
			(0.720)	(0.818)
Percentage of credits after 60-credit threshold attempted at 300 level: > 86%			-5.514**	-1.111
			(0.596)	(0.704)
Took 100-level math course after 60-credit threshold			8.804**	8.004**
			(0.548)	(0.557)
Percentage of credits in first term attempted at 100 level: > 36%			-5.732**	-6.416**
			(0.947)	(0.958)
Observations	5,158	5,158	5,158	5,158
F-test	34.19***	121.26***	119.7***	104.42***
R^2	.062	.261	.271	.299

*$p < .05$. **$p < .01$. ***$p < .001$.

Table 9
Multiple Regression Analysis: State A, Two-Year Entrants, Excess Credits Attempted

Variable	Model 1	Model 2	Model 3	Model 4
Student characteristics				
Pell recipient	-0.263	-2.729*	-1.766	-2.332
	(2.002)	(1.634)	(1.645)	(1.591)
Race/ethnicity				
Unknown	-1.636	-2.045	-0.688	-1.256
	(2.803)	(2.271)	(2.280)	(2.203)
Hispanic	-4.928	-1.674	0.670	0.177
	(5.442)	(4.401)	(4.453)	(4.306)
Black	11.488**	5.790*	5.950*	4.897
	(3.574)	(2.916)	(2.923)	(2.829)
American Indian	7.555	1.869	3.120	3.263
	(13.784)	(11.176)	(11.260)	(10.864)
Asian	5.316	4.403	4.052	3.044
	(6.186)	(4.998)	(5.060)	(4.896)
Female	-5.061**	-0.748	-1.950	-0.478
	(1.574)	(1.305)	(1.295)	(1.275)
Age	-0.239	-0.209	-0.184	-0.194
	(0.170)	(0.138)	(0.139)	(0.134)
Number of developmental education placements	4.818**	5.214**	2.067	2.541*
	(0.806)	(0.658)	(1.264)	(1.227)
Completed associate degree	0.088	-2.535*	0.393	-0.897
	(1.536)	(1.478)	(1.538)	(1.514)
Simple transcript indicators				
Percentage of all credits attempted at a community college		-29.384**		-3.555
		(6.052)		(7.547)
Percentage of credits before 60-credit threshold attempted at 100 level		13.026*		13.074*
		(6.038)		(5.859)
Percentage of credits after 60-credit threshold attempted at 100 level		27.326**		27.233*
		(9.563)		(11.667)
Percentage of credits before 60-credit threshold attempted at 300 level		-82.860**		-72.618**
		(10.016)		(16.339)
Percentage of credits after 60-credit threshold attempted at 300 level		-42.486**		-14.150
		(5.667)		(9.551)
Data-mined cut points and variables				
Percentage of credits after 60-credit threshold attempted at 300 level: > 71%			-15.280**	-8.044**
			(1.635)	(2.232)
Number of developmental education placements: Two or three			7.130**	6.182**
			(2.850)	(2.756)
Percentage of credits after 60-credit threshold earned in STEM: > 4%			6.797**	4.442**
			(1.406)	(1.408)
Percentage of credits after 60-credit threshold earned at 200 level: > 17%			0.586	0.927
			(1.639)	(2.106)
Percentage of credits before 60-credit threshold attempted at 300 level: > 11%			-16.278**	-2.505
			(1.944)	(3.314)
Percentage of all credits earned at a community college: > 45%			-8.553**	-9.649**
			(1.634)	(2.045)
Observations	666	666	666	666
F-test	7.69***	31.53***	29.04***	26.44***
R^2	.105	.421	.417	.463

*$p < .05$. **$p < .01$. ***$p < .001$.

Table 10
Multiple Regression Analysis: State B, Two-Year Entrants, Excess Credits Attempted

Variable	Model 1	Model 2	Model 3	Model 4
Student characteristics				
Pell recipient	3.3312** (0.5809)	3.2300** (0.5483)	2.7685** (0.5302)	3.0400** (0.5216)
Race/ethnicity				
Unknown	1.3641 (1.1841)	1.4269 (1.1081)	0.7158 (1.0783)	1.0143 (1.0542)
Hispanic	1.6033 (1.0999)	1.6741 (1.0296)	0.8744 (1.0014)	0.9614 (0.9796)
Black	12.9728** (0.5758)	10.2080** (0.5449)	8.8169** (0.5315)	8.2124** (0.5217)
Asian	6.4372** (1.1811)	5.5303** (1.1057)	5.0721** (1.0762)	4.6437** (1.0522)
American Indian	9.3327** (2.1122)	6.0241** (1.9868)	7.4878** (1.9244)	6.1777** (1.8905)
Female	-3.1739** (0.3891)	-2.3012** (0.3665)	-1.7965** (0.3560)	-1.5686** (0.3502)
Age	-0.7160** (0.0413)	-0.4551** (0.0399)	-0.4906** (0.0382)	-0.3517** (0.0382)
Number of developmental education placements	5.4488** (0.2304)	7.0398** (0.2818)	4.6063** (0.2278)	6.0866** (0.2703)
Completed associate degree	4.0021** (0.5225)	4.5898** (0.5673)	3.7201** (0.4932)	5.3869** (0.5430)
Simple transcript indicators				
Percentage of all credits attempted at a community college		-27.0656** (1.6426)		-22.6841** (1.7434)
Percentage of credits before 60-credit threshold attempted at 100 level		7.1707** (0.7321)		2.9564** (1.0604)
Percentage of credits after 60-credit threshold attempted at 100 level		29.5217** (1.7256)		24.6551** (1.6497)
Percentage of credits before 60-credit threshold attempted at 300 level		-13.7556** (0.8662)		-10.8325** (0.8393)
Percentage of credits after 60-credit threshold attempted at 300 level		-3.9105** (1.2202)		0.1894 (1.1840)
Percentage of total credits earned in the statewide transfer library		-10.0431** (0.5811)		-1.6245* (0.7608)
Data-mined cut points and variables				
Percentage of credits before 60-credit threshold earned at 100 level: > 44%			5.9398** (0.3864)	2.4593** (0.6271)
Percentage of credits after 60-credit threshold attempted at 200 level: > 10%			9.7003** (0.4653)	7.4236** (0.4775)
Percentage of credits before transfer attempted at 200 level: > 2%			8.7289** (0.8283)	6.6493** (0.8274)
Percentage of credits before transfer attempted at 200 level: 1.8% ~ 2%			28.6430** (0.9386)	25.2486** (0.9438)
Percentage of credits before transfer attempted at 200 level: 1.0% ~ 1.8%			3.2582** (0.4557)	0.5769 (0.5012)
Percentage of total credits earned in the statewide transfer library: > 87%			-8.8657** (0.3570)	-7.2934** (0.4779)
Observations	12,721	12,719	12,721	12,719
F-test	154.15***	234.31***	291.62***	249.44***
R^2	.118	.228	.269	.302

**p* < .05. ***p* < .01. ****p* < .001.

Table 11
Supplemental Regression Analysis Summary

Model Summaries	State A, Four-Year Entrants	State A, Two-Year Entrants	State B, Two-Year Entrants
Model 1 (student characteristics)	R^2 = .06 (F = 34.19, p < .001)	R^2 = .11 (F = 7.69, p < .001)	R^2 = .12 (F = 154.15, p < .001)
Model 2 (student characteristics, simple transcript indicators)	R^2 = .26 (F = 121.26, p < .001)	R^2 = .42 (F = 31.53, p < .001)	R^2 = .23 (F = 234.31, p < .001)
Model 3 (student characteristics, data-mined cut points)	R^2 = .27 (F = 119.7, p < .001)	R^2 = .42 (F = 29.04, p < .001)	R^2 = .27 (F = 291.62, p < .001)
Model 4 (student characteristics, simple transcript indicators, data-mined cut points)	R^2 = .30 (F = 104.42, p < .001)	R^2 = .46 (F = 26.44, p < .001)	R^2 = .30 (F = 249.44, p < .001)
Significant Indicators, Model 4[1]			
Student characteristics (beta)	• Pell recipient (1.82) • Black (5.97) • Female (-1.87)		• Pell recipient (3.04) • Black (8.21) • Asian (4.64) • American Indian (6.18) • Female (-1.57) • Age (-0.35) • Number of developmental education placements (6.09) • Completed associate degree (5.39)
Simple transcript indicators (beta)	• Percentage of all credits attempted at a community college (31.28) • Percentage of credits before 60-credit threshold attempted at 300 level (-23.20) • Percentage of credits after 60-credit threshold attempted at 300 level (-36.10)	• Percentage of credits before 60-credit threshold attempted at 300 level (-72.62)	• Percentage of all credits attempted at a community college (-22.68) • Percentage of credits before 60-credit threshold attempted at 100 level (2.96) • Percentage of credits after 60-credit threshold attempted at 100 level (24.66) • Percentage of credits before 60-credit threshold attempted at 300 level (-10.83)
Data-mined cut points (beta)	• Took 100-level math course after 60-credit threshold (8.00) • Percentage of credits in first term attempted at 100 level: > 36% (-6.42)	• Percentage of credits after 60-credit threshold attempted at 300 level: > 71% (-8.04) • Number of developmental education placements: 2 or 3 (6.18) • Percentage of credits after 60-credit threshold earned in STEM: > 4% (4.44) • Percentage of all credits earned at a community college: > 45% (-9.65)	• Percentage of credits before 60-credit threshold earned at 100 level: > 44% (2.46) • Percentage of credits after 60-credit threshold attempted at 200 level: > 10% (7.42) • Percentage of credits before transfer attempted at 200 level: > 2% (6.65) • Percentage of credits before transfer attempted at 200 level: 1.8% ~ 2% (25.25) • Percentage of total credits earned in the statewide transfer library: > 87% (-7.29)

[1] Model 4 independent variables significant at p < .01.

4. Discussion

In general, we found much overlap between states and between two- and four-year entrants in the variables associated with higher or lower numbers of excess credits attempted among bachelor's graduates. Across the samples, students who earned a bachelor's degree with fewer excess credits took fewer 100-level courses overall and, more specifically, took more 300-level courses, fewer 200-level courses, and no 100-level math courses after accumulating 60 college-level credits. Some students in State A who graduated with more excess credits also earned larger proportions of STEM credits after reaching the 60-credit threshold, and took more 100-level courses and fewer 300-level courses in any subject immediately after transferring into the four-year institution. Unfortunately, detail on STEM courses was not available from State B. Among two-year entrants, other indicators of community college course-taking were associated with more excess credits. For example, two-year entrants with more excess credits had more developmental education placements and took larger proportions of credits at a community college (State A) and took more 100-level courses before 60 credits and fewer courses in the statewide transfer library (State B).

The supplemental multiple regression analysis compared the relative effects of student characteristics and course-taking behaviors on students' excess credits. Overall, the regression models that included student course-taking behavior variables explained between two and four times the amount of variance in excess credits among bachelor's earners compared with the models that only included student characteristics (e.g., demographics, developmental education placement, completion of an associate degree). Comparing Models 2 and 3 across the samples, there does not seem to be a substantial difference in the overall explanatory power with the inclusion of relatively simple transcript variables or more complex transcript variables identified through data-mining analysis. Our results suggest that the percentage of 100- and 300-level credits students attempt before and after accumulating 60 credits, as well as the overall percentage of credits they attempt at a community college, have similar power to explain variation of students' excess credits as more complex transcript indicators identified using partition trees. However, it is important to note that the simpler course-taking variables reflect enrollment patterns identified through the data-mining analysis—patterns that we likely would not have identified without the data mining.

The models reveal other notable findings about the effects of course-taking patterns on excess credits for particular student demographic groups. In the first model only including student characteristics among two-year entrants in State A, we found significant associations between attempting more excess credits and being Black, male, or placed into developmental education. However, these associations were no longer statistically significant when later models included course-taking behaviors as additional independent variables. A similar pattern for student race/ethnicity and number of developmental education placements was observed among two-year entrants in State B and among four-year entrants in State A, though the effects remained significant. (Note the sample size was much larger in State B.) These findings suggest that the barriers to efficient transfer for students of color and for academically underprepared students could be at least partially mitigated if colleges made it easier for these students to progress into and through academic programs, thereby sequencing coursework from lower to upper division as they progressed toward completion.

4.1 Implications for Practice

Many two- and four-year entrants in our study were not able to complete 100- and 200-level courses and move on to 300- and 400-level courses, often their major-specific courses, by the time they earned 60 credits. This finding suggests that students could minimize their excess credits at graduation by exploring fields of interest and deciding on a major field early on, so that by the time they accumulate 60 credits they will have taken the right lower division major-prerequisite courses to be able to take major-specific 300- and 400-level courses. For leaders at community colleges and four-year institutions, our findings highlight the importance of early advising and other supports focused on helping students explore career and academic options and choose a program of study. This is often not a major focus of community college advising. Rather, community college advisors often encourage students seeking to transfer to "get their general education requirements" out of the way, on the assumption that doing so will give them the greatest flexibility to choose a major when they transfer (Bailey, Jaggars, & Jenkins, 2015, pp. 27–31). The findings from this analysis suggest that this may be bad advice. To avoid earning excess credits, rather than taking just any lower division courses that meet general education distribution requirements, students should take 100- and 200-level courses that are required for their major field of interest so that they can take major-specific upper division courses by the time they reach 60 credits. Additionally, advising students to explore and choose a major-specific pathway early on may help them avoid taking too many 100-level courses after they have accumulated 60 credits, a pattern which we found to be prevalent among students with more excess credits.

Our findings reinforce Wyner et al.'s (2016) recommendation that two- and four-year institutions work together to create clear programmatic transfer maps and guidelines so that students take the right lower division and pre-major coursework and thus minimize inefficient credit transfer. In their *Transfer Playbook*, Wyner et al. found that the high-performing partnerships of two- and four-year institutions they identified tended to work together to clearly map out field- and major-specific transfer pathways and to focus advising on helping students choose a program of study early on and ensuring that students take courses that will apply toward a bachelor's degree in their desired major. That we found similar course-taking indicators of excess credits among four-year entrants suggests that the benefits of clearly defined degree pathways and focused advising are not unique to community college transfer students. Indeed, hundreds of colleges and universities across the country are undertaking "guided pathways" reforms to better map course sequences and progress milestones to increase completion rates and decrease excess credits and time to degree for all of their students. While transfer students may encounter more difficulty navigating the complex pathway to a bachelor's degree across multiple institutions, our findings suggest that clarifying paths to degrees and supporting students in choosing a program direction early on can lead to more efficient degree completion for community college and university entrants alike.

4.2 Limitations

We measured credit transfer efficiency using the number of excess credits among bachelor's degree completers. Examining excess credits among completers allows for a full view of how efficiently students progress to completion, including the efficiency of the credit transfer process. However, measuring credit transfer efficiency using excess credits among bachelor's degree completers omits those students who do not complete a bachelor's degree. This limitation is particularly notable given that students who experience credit loss at transfer are less likely to complete a bachelor's degree (Monaghan & Attewell, 2015). Focusing on excess credits among completers, as we do in the current study, likely understates the consequences of credit transfer efficiency by not taking into account how inefficient credit transfer lowers students' chances of completing a bachelor's degree.

4.3 Directions for Future Research

Given the exploratory nature of this study, there are a number of future directions for research based on our findings. The data-mining techniques used here allowed us to explore types of course-taking behaviors that might be important in explaining excess credits. Institutional researchers at community colleges and universities could replicate our analyses to better understand course-taking behaviors and other factors associated with excess credits among completers at their institutions. Although data-mining techniques are not yet frequently used in higher education research, the information presented here can inform more traditional analytic methods. Importantly, using data-mining techniques to identify prevalent patterns in existing data can make researchers aware of relevant explanatory variables that may have been overlooked by theory, and as a result, by previous research efforts. Translating these findings into more sophisticated models, which consider each of the variables independently and introduce the interactions identified in the partitioning analysis, can allow for a more complete understanding of the most relevant factors in explaining the phenomenon of interest across diverse student populations (Attewell & Monaghan, 2015). This approach is particularly useful when attempting to build models that test topics with a limited theoretical or empirical understanding, such as the relationship between course-taking behaviors and transfer efficiency.

4.4 Conclusion

This study used data-mining techniques to explore course-taking behaviors among two- and four-year college entrants who earned bachelor's degrees in order to better understand transfer credit efficiency as measured by excess credits. Regardless of whether students entered at or transferred to a four-year college, those who took 100- and 200-level courses that enabled them to take mostly 300- and 400-level courses after they reached 60 credits along their path to a bachelor's degree graduated with fewer excess credits on average. These findings emphasize the importance for colleges and universities of working both internally and with their transfer partners to create clear programmatic degree pathways, to help students to explore and select a major field early on, and to ensure that students continue to take courses that will apply to a degree in their intended major as they progress.

References

Adelman, C. (2005). *Moving into town—and moving on. The community college in the lives of traditional-age students*. Washington, DC: U.S. Department of Education.

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. Washington, DC: U.S. Department of Education.

Attewell, P., & Monaghan, D. (2015). *Data mining for social sciences: An introduction.* Oakland, CA: University of California Press.

Bahr, P. R. (2013). The deconstructive approach to understanding community college students' pathways and outcomes. *Community College Review*, *41*(2), 137–153.

Bailey, T. R., Jaggars, S. S., & Jenkins D. (2015). *Redesigning America's community colleges: A clearer path to student success.* Cambridge, MA: Harvard University Press.

Baker, R. (2016). The effects of structured transfer pathways in community colleges. *Educational Evaluation and Policy Analysis*, *38*(4), 626–646.

Belfield, C., Crosta, P., & Jenkins, D. (2014). Can community colleges afford to improve completion? Measuring the costs and efficiency effects of college reforms. *Educational Evaluation and Policy Analysis*, *36*(3), 327–346.

Belfield, C., Fink, J., & Jenkins, D. (2017). *Is it really cheaper to start at a community college? The consequences of inefficient transfer for community college students seeking bachelor's degrees* (CCRC Working Paper No. 94). New York, NY: Columbia University, Teachers College, Community College Research Center.

Belfield, C., Jenkins, D., & Lahr, H. (2016). *Momentum: The academic and economic value of a 15-credit first-semester course load for college students in Tennessee* (CCRC Working Paper No. 88). New York, NY: Columbia University, Teachers College, Community College Research Center.

Bragg, D. D. (2012). Two-year college mathematics and student progression in STEM programs of study. In S. Olson & J. B. Labov (Eds.), *Community colleges in the evolving STEM education landscape* (pp. 80–101). Washington, DC: The National Academies Press.

Calcagno, J. C., Crosta, P., Bailey, T., & Jenkins, D. (2007). Stepping stones to a degree: The impact of enrollment pathways and milestones on community college students. *Research in Higher Education*, *48*(7), 775–801.

Complete College America. (2011). *Time is the enemy.* Washington, DC: Author. Retrieved from https://completecollege.org/wp-content/uploads/2017/08/Time_Is_the_Enemy.pdf

Complete College America. (n.d.). *Data dashboard*. Retrieved from https://completecollege.org/data-dashboard/

Cullinane, J. P. (2014). *The path to timely completion: Supply- and demand-side analyses of time to bachelor's degree completion* (Doctoral dissertation). Retrieved from https://repositories.lib.utexas.edu/handle/2152/24932

Fink, J. (2017). *Visualizing the many routes community college students take to complete a bachelor's degree*. Retrieved from https://ccrc.tc.columbia.edu/blog/visualizing-many-routes-bachelors-degree.html

Fink, J. & Jenkins, D. (2017). Takes two to tango: Essential practices of highly effective transfer partnerships. *Community College Review*, *45*(4), 294–310.

Hagedorn, L. S. (2005). Transcript analyses as a tool to understand community college student academic behaviors. *Journal of Applied Research in the Community College*, *13*(1), 45–57.

Hagedorn, L. S., & Kress, A. M. (2008). Using transcripts in analyses: Directions and opportunities. *New Directions for Community Colleges*, *2008*(143), 7–17.

Handel, S. J., & Williams, R. A. (2012). *The promise of the transfer pathway: Opportunity and challenge for community college students seeking the baccalaureate degree*. New York, NY: College Board.

Hodara, M., Martinez-Wenzl, M., Stevens, D., & Mazzeo, C. (2016). *Improving credit mobility for community college transfer students: Findings and recommendations from a 10-state study.* Portland, OR: Education Northwest.

Jaggars, S. S., Fink, J., Fletcher, J., & Dundar, A. (2016). *A longitudinal analysis of community college pathways to computer science bachelor's degrees.* Mountain View, CA: Google Inc.

Jenkins, D., & Fink, J. (2016). *Tracking transfer: New measures of state and institutional effectiveness in helping community college students attain bachelor's degrees.* New York, NY: Community College Research Center, Aspen Institute, and the National Student Clearinghouse Research Center.

Leinbach, D. T., & Jenkins, D. (2008). *Using longitudinal data to increase community college student success: A guide to measuring milestones and momentum point attainment* (CCRC Research Tools No. 2). New York, NY: Columbia University, Teachers College, Community College Research Center.

Monaghan, D., & Attewell, P. (2015). The community college route to the bachelor's degree. *Educational Evaluation and Policy Analysis*, *37*(1), 70–91.

Simone, S. A. (2014). *Transferability of postsecondary credit following student transfer or coenrollment* (NCES 2014-163). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

Texas Higher Education Coordinating Board. (2001). *Transfer issues advisory committee report: Identifying and closing the gaps*. Austin, TX: Author.

U.S. Government Accountability Office. (2017). *Students need more information to help reduce challenges in transferring college credits.* Washington, DC: Author.

Wang, X. (2016). Course-taking patterns of community college students beginning in STEM: Using data mining techniques to reveal viable STEM transfer pathways, *Research in Higher Education*, *57*(5), 544–569.

Washington State Higher Education Coordinating Board. (2011). *Transfer report.* Olympia, WA: Author.

Wyner, J., Deane, K. C., Jenkins, D., & Fink, J. (2016). *The transfer playbook: Essential practices for two- and four-year institutions.* New York, NY: The Aspen Institute College Excellence Program.

Xu, D., Jaggars, S. S., & Fletcher, J. (2016). *How and why does two-year college entry influence baccalaureate aspirants' academic and labor market outcomes?* (CAPSEE Working Paper). New York, NY: Center for Analysis of Postsecondary Education and Employment.

Appendix A: Figures

Figure A1
Partition Tree: State A, Four-Year Entrants, Excess Credits Attempted

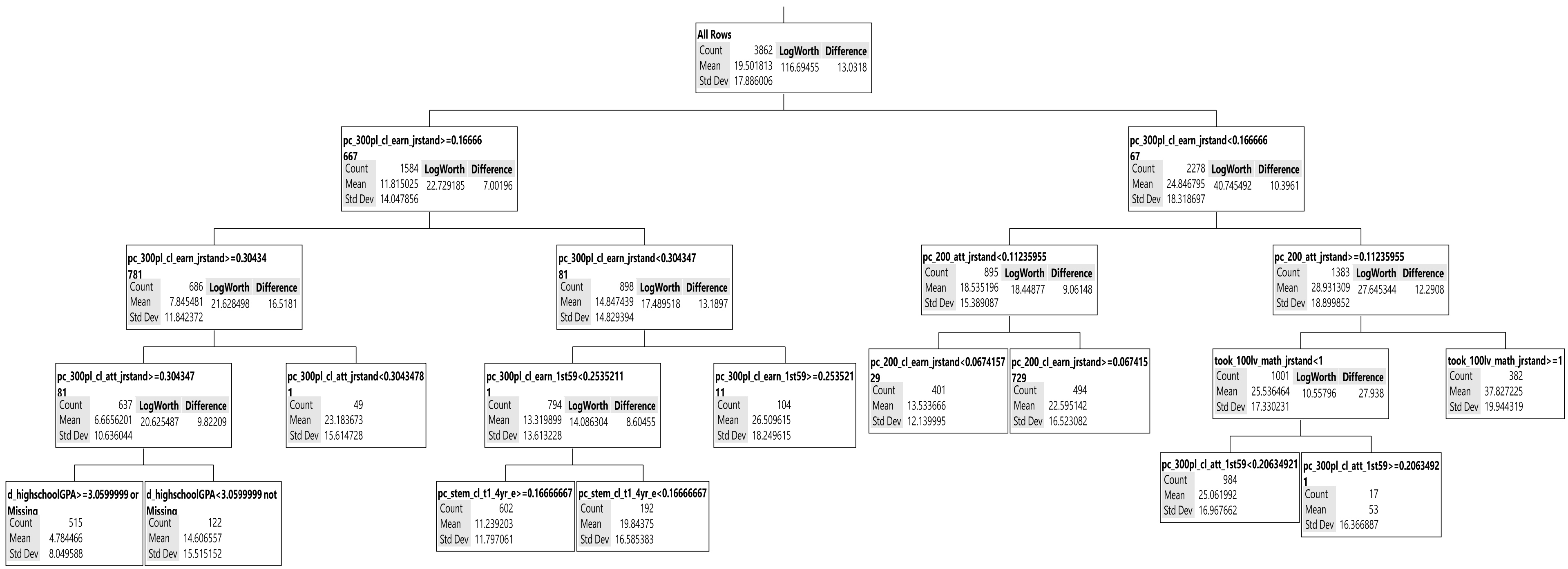


Figure A2
Partition Tree: State A, Two-Year Entrants, Excess Credits Attempted

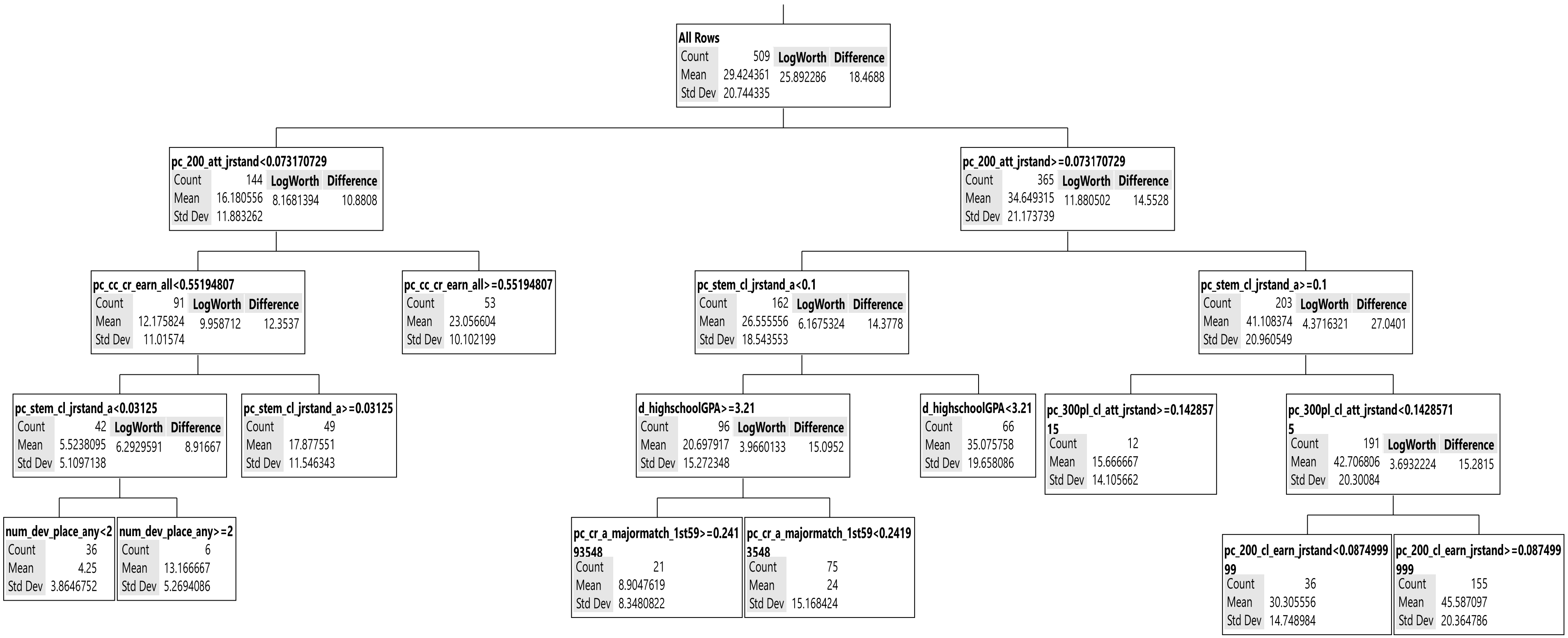
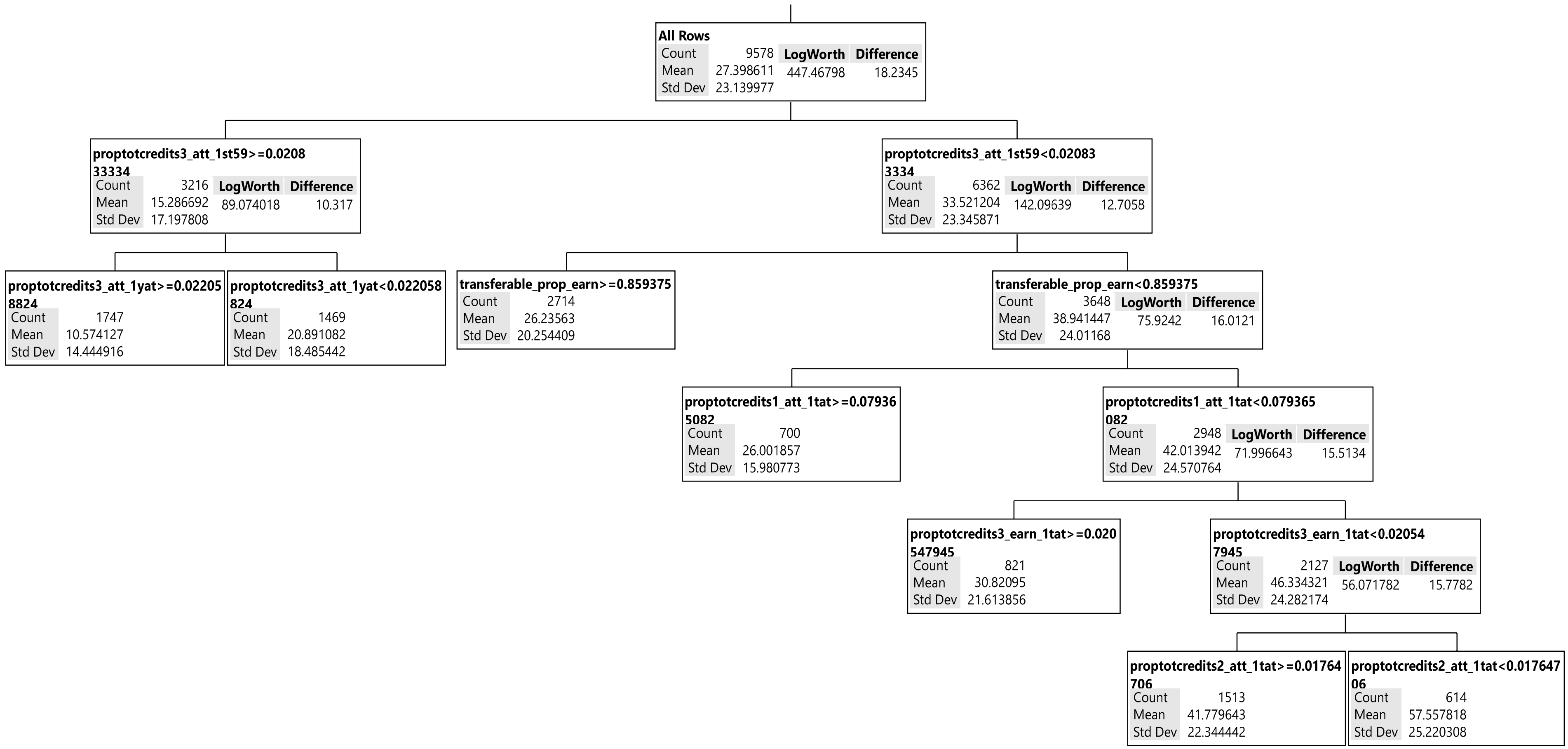


Figure A3
Partition Tree: State B, Two-Year Entrants, Excess Credits Attempted



Appendix B: Tables From Analyses of Excess Credits Earned

Table B1
Model Summary and Column Contributions:
State A, Four-Year Entrants, Excess Credits Earned

	R^2	*n*	Result of First Split on Excess Credits[1]	Number of Splits	Portion
Training	.565	3,905		57	
Validation	.458	1,253			
Variable					
Percentage of credits attempted at 200 level after 60 credits			More excess	1	0.3031
Percentage of credits earned at 200 level after 60 credits			++	8	0.127
Percentage of credits earned at 300+ level before 60 credits			More excess	7	0.125
Percentage of college-level credits earned at 300+ level after 60 credits			Fewer excess	17	0.1172
Percentage of credits attempted at 200 level after 60 credits			More excess	1	0.0878
Percentage of college-level credits earned at 300+ level before 60 credits			More excess	11	0.08
Percentage of college-level credits attempted at 300+ level before 60 credits			More excess	1	0.0521
Attempted 100-level math after 60 credits			More excess	2	0.0472
Percentage of college-level credits earned at 200 level after 60 credits			++	1	0.0197
Percentage of credits earned in major area before 60 credits			++	1	0.0159
Percentage of credits earned in major area one year after transfer			++	1	0.0139
Percentage of college-level credits earned in STEM after 60 credits			++	1	0.0042
Percentage of college-level credits attempted in STEM before 60 credits			++	1	0.0038
Percentage of college-level credits earned at 200 level before 60 credits			++	1	0.0009
Percentage of credits attempted at 300+ level before 60 credits			++	1	0.0009
Percentage of credits attempted in major area one term after transfer			++	1	0.0007
Percentage of college-level credits attempted at 300+ level after 60 credits			Fewer excess	1	0.0003

[1] This column describes the directionality of the first time a given variable was used to partition the data. Higher values (e.g., larger percentages, higher GPAs) had either more or fewer excess credits on average, as indicated in this column.

++ Variables used to partition data beyond the eighth split are not visually examined.

Table B2
Model Summary and Column Contributions:
State A, Two-Year Entrants, Excess Credits Earned

	R^2	*n*	Result of First Split on Excess Credits[1]	Number of Splits	Portion
Training	.537	490		17	
Validation	.283	176			
Variable					
Percentage of credits earned at 200 level after 60 credits			More excess	1	0.3625
Percentage of credits attempted at 300+ level after 60 credits			Fewer excess	2	0.1433
Percentage of credits attempted at the community college			Fewer excess	2	0.1266
Percentage of credits attempted at 200 level after 60 credits			More excess	1	0.0781
Percentage of credits earned in STEM before 60 credits			More excess	2	0.0606
Completed general education requirement: Science			More excess	1	0.0591
Completed certificate			More excess	1	0.0493
Number of developmental education placement areas			More excess	1	0.0331
Percentage of credits earned in STEM after 60 credits			More excess	1	0.0245
Percentage of credits earned at 300+ level after 60 credits			Fewer excess	1	0.0208
Attempted 100-level math after 60 credits			More excess	1	0.0184
Percentage of credits earned at 200 level one term after transfer			Fewer excess	1	0.0162
Percentage of credits earned at 100 level before 60 credits			More excess	1	0.0054
Completed general education requirement: Communications			Fewer excess	1	0.002

[1] This column describes the directionality of the first time a given variable was used to partition the data. Higher values (e.g., larger percentages, higher GPAs) had either more or fewer excess credits on average, as indicated in this column.

++ Variables used to partition data beyond the eighth split are not visually examined.

Table B3
Model Summary and Column Contributions:
State B, Two-Year Entrants, Excess Credits Earned

	R^2	n	Result of First Split on Excess Credits[1]	Number of Splits	Portion
Training	.588	9,444		116	
Validation	.516	3,278			
Variable					
Percentage of credits attempted in 100-level courses before 60 credits			More excess	1	0.3213
Percentage of credits earned in 300-level courses one term after transfer			Fewer excess	13	0.1498
Percentage of credits earned in 300-level courses before 60 credits			Fewer excess	16	0.1477
Percentage of credits earned in 200-level courses one term after transfer			++	30	0.0910
Percentage of credits earned at the community college			++	8	0.0684
Percentage of credits earned in transferable courses			++	4	0.0583
Percentage of credits attempted in 100-level courses one term after transfer			++	5	0.0287
Percentage of credits attempted in 200-level courses after 60 credits			More excess	3	0.0224
Percentage of credits attempted in 300-level courses after 60 credits			++	3	0.0197
Percentage of credits earned in 200-level courses after 60 credits			++	4	0.0172
Percentage of credits earned in 100-level courses one term after transfer			++	3	0.0129
Percentage of credits earned in 100-level courses after 60 credits			++	5	0.0127
Percentage of credits earned in 300-level courses after 60 credits			++	3	0.0125
Percentage of credits earned in 200-level courses before 60 credits			++	6	0.0117
Percentage of credits attempted in major area			++	1	0.0087
Percentage of credits earned in 100-level courses before 60 credits			++	2	0.0069
Age of first enrollment			++	3	0.0032
Four-digit CIP code for major			++	2	0.0025
Percentage of credits earned in 300-level courses one year after transfer			++	1	0.0020
Completed general education requirement: Social science			++	1	0.0019
Took a 100-level English course one year after transfer			++	1	0.0004
Percentage of credits attempted at the community college			++	1	0.0001

[1] This column describes the directionality of the first time a given variable was used to partition the data. Higher values (e.g., larger percentages, higher GPAs) had either more or fewer excess credits on average, as indicated in this column.

++ Variables used to partition data beyond the eighth split are not visually examined.

Appendix C: List of Variable Names by State

Student Characteristic	Variable Name	
	State A	State B
Demographics		
High school GPA	d_highschoolGPA	-
Pell grant recipient	d_pell	-
Received any financial aid	-	receive_finaid
White	d_race_white	race_w
Unknown race/ethnicity	d_race_unknwn	race_unknown
Native Pacific Islander	d_race_nativepacific	-
Multiracial	d_race_multiracial	-
Hispanic	d_race_hispanic	race_h
Black	d_race_black	race_b
Asian	d_race_asian	race_a
Native American	d_race_americanindian	race_ai
Native Alaskan	d_race_alaskanative	-
Other race/ethnicity	-	race_oth
Male	d_male	-
Female	d_female	female
Age at enrollment	d_enroll_age	enroll_age
Awards		
Completed any associate degree	AA_completer	any_aa_bt
Completed associate of applied science	-	aas_bt
Completed associate of science	-	as_bt
Completed certificate	cert_completer	-
Earned any award at community college (two-year entrants only)	ever_pretransfer_awd	-
Developmental education placement		-
Number of developmental education placement areas	num_dev_place_any	-
Number of developmental education placement areas taken at the community college	num_dev_place_cc	-

	Variable Name	
Student Characteristic	State A	State B
Gatekeeper courses		
Attempted 100-level English one term after transfer (two-year entrants only)	took_100lv_eng_t1_4yr	eng1_1tat
Attempted 100-level math one term after transfer (two-year entrants only)	took_100lv_math_t1_4yr	math1_1tat
Attempted 100-level English one year after transfer (two-year entrants only)	took_100lv_eng_t2nosum	eng1_1yat
Attempted 100-level math one year after transfer (two-year entrants only)	took_100lv_math_t2nosum	math1_1yat
Attempted 100-level English before 60 college-level credit threshold	took_100lv_eng_1st59	eng1_1st59
Attempted 100-level math before 60 college-level credit threshold	took_100lv_math_1st59	math1_1st59
Attempted 100-level English after 60 college-level credit threshold	took_100lv_eng_jrstand	eng1_jr
Attempted 100-level math after 60 college-level credit threshold	took_100lv_math_jrstand	math1_jr
Community college course-taking		
Percentage of credits attempted at the community college	pc_cc_cr_att_all	propcc_att
Percentage of college-level credits attempted at the community college	pc_cc_cr_att_cl	-
Percentage of credits earned at the community college	pc_cc_cr_earn_all	propcc_earn
Percentage of college-level credits earned at the community college	pc_cc_cr_earn_cl	-
Percentage of credits attempted at the community college taken in the transfer course library		transferable_prop_att
Percentage of credits earned at the community college taken in the transfer course library		transferable_prop_earn
General education courses		-
Completed general education requirement: Communications (two-year entrants only)	com_complete	-
Completed general education requirement: Humanities (two-year entrants only)	hum_complete	hum_ge_12
Took a general education course in humanities	-	hum_12
Completed general education requirement: Social science (two-year entrants only)	soc_complete	soc_ge_12
Took a general education course in social science	-	soc_ge
Completed general education requirement: History (two-year entrants only)	hist_complete	-
Completed general education requirement: Science (two-year entrants only)	sci_complete	sci_ge_8
Took a general education course in math	-	math_6
Completed general education requirement: Math (two-year entrants only)	math_complete	math_ge_6
Completed general education requirements: All (two-year entrants only)	gened_complete	-

	Variable Name	
Student Characteristic	State A	State B
Course-taking by course level		
Percentage of credits attempted at 100 level pre-transfer (two-year entrants only)	pc_100_cl_att_prexfer	propcredits1_cc_att
Percentage of credits attempted at 200 level pre-transfer (two-year entrants only)	pc_200_cl_att_prexfer	propcredits2_cc_att
Percentage of credits attempted at 300+ level pre-transfer (two-year entrants only)	pc_300pl_cl_att_prexfer	-
Percentage of credits attempted at 100 level one term after transfer (two-year entrants only)	pc_100_cl_att_t1	proptotcredits_att_1tat
Percentage of credits attempted at 200 level one term after transfer (two-year entrants only)	pc_200_cl_att_t1	proptotcredits2_att_1tat
Percentage of credits attempted at 300+ level one term after transfer (two-year entrants only)	pc_300pl_cl_att_t1	proptotcredits3_att_1tat
Percentage of credits attempted at 100 level one year after transfer (two-year entrants only)	pc_100_cl_att_t2nosum	proptotcredits_att_1yat
Percentage of credits attempted at 200 level one year after transfer (two-year entrants only)	pc_200_cl_att_t2nosum	proptotcredits2_att_1yat
Percentage of credits attempted at 300+ level one year after transfer (two-year entrants only)	pc_300pl_cl_att_t2nosum	proptotcredits3_att_1yat
Percentage of credits attempted at 100 level before 60 college-level credit threshold	pc_100_cl_att_1st59	proptotcredits1_att_jr
Percentage of credits attempted at 200 level before 60 college-level credit threshold	pc_200_cl_att_1st59	proptotcredits2_att_jr
Percentage of credits attempted at 300+ level before 60 college-level credit threshold	pc_300pl_cl_att_1st59	proptotcredits3_att_jr
Percentage of credits attempted at 100 level after 60 college-level credit threshold	pc_100_cl_att_jrstand	proptotcredits1_att_1st59
Percentage of credits attempted at 200 level after 60 college-level credit threshold	pc_200_cl_att_jrstand	proptotcredits2_att_1st59
Percentage of credits attempted at 300+ level after 60 college-level credit threshold	pc_300pl_cl_att_jrstand	proptotcredits3_att_1st59
Percentage of credits earned at 100 level pre-transfer (two-year entrants only)	pc_100_cl_earn_prexfer	propcredits1_cc_earn
Percentage of credits earned at 200 level pre-transfer (two-year entrants only)	pc_200_cl_earn_prexfer	propcredits2_cc_earn
Percentage of credits earned at 300+ level pre-transfer (two-year entrants only)	pc_300pl_cl_earn_prexfer	-
Percentage of credits earned at 100 level one term after transfer (two-year entrants only)	pc_100_cl_earn_t1	proptotcredits_earn_1tat
Percentage of credits earned at 200 level one term after transfer (two-year entrants only)	pc_200_cl_earn_t1	proptotcredits2_earn_1tat
Percentage of credits earned at 300+ level one term after transfer (two-year entrants only)	pc_300pl_cl_earn_t1	proptotcredits3_earn_1tat
Percentage of credits earned at 100 level one year after transfer (two-year entrants only)	pc_100_cl_earn_t2nosum	proptotcredits_earn_1yat
Percentage of credits earned at 200 level one year after transfer (two-year entrants only)	pc_200_cl_earn_t2nosum	proptotcredits2_earn_1yat
Percentage of credits earned at 300+ level one year after transfer (two-year entrants only)	pc_300pl_cl_earn_t2nosum	proptotcredits3_earn_1yat
Percentage of credits earned at 100 level before 60 college-level credit threshold	pc_100_cl_earn_1st59	proptotcredits1_earn_jr
Percentage of credits earned at 200 level before 60 college-level credit threshold	pc_200_cl_earn_1st59	proptotcredits2_earn_jr
Percentage of credits earned at 300+ level before 60 college-level credit threshold	pc_300pl_cl_earn_1st59	proptotcredits3_earn_jr
Percentage of credits earned at 100 level after 60 college-level credit threshold	pc_100_cl_earn_jrstand	proptotcredits1_earn_1st59
Percentage of credits earned at 200 level after 60 college-level credit threshold	pc_200_cl_earn_jrstand	proptotcredits2_earn_1st59
Percentage of credits earned at 300+ level after 60 college-level credit threshold	pc_300pl_cl_earn_jrstand	proptotcredits3_earn_1st59

Student Characteristic	Variable Name: State A	Variable Name: State B
Within-major course-taking		
Code of bachelor’s degree	final_student_major_2cip	major_cip_4
Percentage of credits attempted in major area	-	propbtcredits_degree_att
Percentage of credits earned in major area at the community college	pc_cr_e_majormatch_CC	-
Percentage of credits earned in major area one term after transfer (two-year entrants only)	pc_cr_e_majormatch_t1	-
Percentage of credits earned in major area one year after transfer (two-year entrants only)	pc_cr_e_majormatch_t2nosum	-
Percentage of credits earned in major area before 60 college-level credit threshold	pc_cr_e_majormatch_1st59	-
Percentage of credits earned in major area after 60 college-level credit threshold	pc_cr_e_majormatch_jrstand	-
Percentage of credits earned in major area pre-transfer (two-year entrants only)	pc_cr_e_majormatch_prexfer	-
Percentage of credits attempted in major area at the community college	pc_cr_a_majormatch_CC	-
Percentage of credits attempted in major area one term after transfer (two-year entrants only)	pc_cr_a_majormatch_t1	-
Percentage of credits attempted in major area one year after transfer (two-year entrants only)	pc_cr_a_majormatch_t2nosum	-
Percentage of credits attempted in major area before 60 college-level credit threshold	pc_cr_a_majormatch_1st59	-
Percentage of credits attempted in major area after 60 college-level credit threshold	pc_cr_a_majormatch_jrstand	-
Percentage of credits attempted in major area pre-transfer (two-year entrants only)	pc_cr_a_majormatch_prexfer	-
Percentage of college-level credits earned in major area at the community college	pc_cl_cr_e_majormatch_CC	-
Percentage of college-level credits earned in major area one term after transfer (two-year entrants only)	pc_cl_cr_e_majormatch_t1	-
Percentage of college-level credits earned in major area one year after transfer (two-year entrants only)	pc_cl_cr_e_majormatch_t2nosum	-
Percentage of college-level credits earned in major area before 60 college-level credit threshold	pc_cl_cr_e_majormatch_1st59	-
Percentage of college-level credits earned in major area after 60 college-level credit threshold	pc_cl_cr_e_majormatch_jrstand	-
Percentage of college-level credits earned in major area pre-transfer (two-year entrants only)	pc_cl_cr_e_majormatch_prexfer	-
Percentage of college-level credits attempted in major area at the community college	pc_cl_cr_a_majormatch_CC	-
Percentage of college-level credits attempted in major area one term after transfer (two-year entrants only)	pc_cl_cr_a_majormatch_t1	-
Percentage of college-level credits attempted in major area one year after transfer (two-year entrants only)	pc_cl_cr_a_majormatch_t2nosum	-
Percentage of college-level credits attempted in major area before 60 college-level credit threshold	pc_cl_cr_a_majormatch_1st59	-
Percentage of college-level credits attempted in major area after 60 college-level credit threshold	pc_cl_cr_a_majormatch_jrstand	-
Percentage of college-level credits attempted in major area pre-transfer (two-year entrants only)	pc_cl_cr_a_majormatch_prexfer	-

	Variable Name	
Student Characteristic	State A	State B
STEM course-taking		
Percentage of college-level credits attempted in STEM at the community college	pc_stem_cl_CC_a	-
Percentage of college-level credits earned in STEM at the community college	pc_stem_cl_CC_e	-
Percentage of college-level credits attempted in STEM pre-transfer (two-year entrants only)	pc_stem_cl_CC_prexfer_a	-
Percentage of college-level credits earned in STEM at pre-transfer (two-year entrants only)	pc_stem_cl_CC_prexfer_e	-
Percentage of college-level credits attempted in STEM at the four-year college	pc_stem_cl_4yr_a	-
Percentage of college-level credits earned in STEM at the four-year college	pc_stem_cl_4yr_e	-
Percentage of college-level credits attempted in STEM one term after transfer (two-year entrants only)	pc_stem_cl_t1_4yr_a	-
Percentage of college-level credits earned in STEM one term after transfer (two-year entrants only)	pc_stem_cl_t1_4yr_e	-
Percentage of college-level credits attempted in STEM one year after transfer (two-year entrants only)	pc_stem_cl_t2nosum_a	-
Percentage of college-level credits earned in STEM one year after transfer (two-year entrants only)	pc_stem_cl_t2nosum_e	-
Percentage of college-level credits attempted in STEM before 60 college-level credit threshold	pc_stem_cl_1st59_a	-
Percentage of college-level credits earned in STEM before 60 college-level credit threshold	pc_stem_cl_1st59_e	-
Percentage of college-level credits attempted in STEM after 60 college-level credit threshold	pc_stem_cl_jrstand_a	-
Percentage of college-level credits earned in STEM after 60 college-level credit threshold	pc_stem_cl_jrstand_e	-