

December 2024

PDP Insights

Early Indicators of Timely Completers



Table of Contents

About this Report.....	2
Highlights	3
Section 1: Descriptive Analysis of Early Success Indicators by Completion Status	4
Section 2: Mixed Effects Logistic Regression Analysis.....	5
Section 3: Special Analysis of Bachelor’s Earning Associate/Certificate Seekers.....	9
Section 4: Summary	9
Methodological Notes.....	11

Early Indicators of Timely Completers

About this Report

The purpose of this report is to analyze the differences between timely completers and students who did not complete in either a timely manner or complete at all, on several key early success indicators that are outlined below. Timely completion has been shown to mitigate financial burdens associated with prolonged enrollment, optimize institutional resource utilization, and enhance the socioeconomic mobility of students, particularly those from underrepresented groups (Bound et al., 2012, 2015; Ortagus et al., 2021; Al-Haddad et al., 2018). By investigating early success indicators such as first-year GPA, credit completion ratios, credits earned, and second-year enrollment, this research seeks to understand the impact of early indicators of student success on timely completion across varying credential types and demographic profiles. These predictors have been analyzed through descriptive statistics and a regression model.

Timely completion is defined, in this report, as the student having earned the credential they initially sought, at any institution, within a specific time frame. For this report, completion at 100%, 150%, and 200% time were included. As timely completion varies depending on the credential sought, students were identified as either bachelor seeking or associate/certificate seeking (see Methodological Notes for more information).

The following early success indicators were analyzed for both groups of credential-seeking students:

- First-Year Credit Completion Ratio (CCR): the ratio of credits earned to credits attempted in a student's first year of enrollment.
- First-Year Credits Earned: the total credit hours a student earned (passing the course) during their first year, regardless of the number attempted.
- First-Year Grade Point Average (GPA): a student's GPA at the end of the first academic year based on the course data provided by the institutions.
- Second-Year Enrollment: a student is considered to have second-year enrollment if they completed a credential by the second year of enrollment or have continued enrollment into their second academic year, either at the same or different institution (retention and transfer, respectively).

The cohort examined in this report consists of credential-seeking (associate/certificate or bachelor's degree), full-time starters who first enrolled during the fall semester of the 2016-17 academic year. To be included in the analysis, each student must have complete information reported on all early success indicators and completion status (see Methodological Notes for more information). The final sample consists of 307,500 unique students—63% were Bachelor's seeking students (194,400 students from 141 institutions) and 37% were associate/certificate seeking students (113,100 students from 191 institutions).

It is important to note that institutions actively opt in to the Postsecondary Data Partnership (PDP). Therefore, while this report is able to uniquely characterize the early success indicators and completion status for a large sample of students, this is not a nationally representative set of institutions. Institutions that participate may vary from others in their student outcomes, demographics, programs, and services. Findings in this report should not be considered representative of the national population of students.

Highlights

The standard time frames typically associated with program completion for full-time students are four years for a bachelor's degree and two years for an associate or certificate program (i.e., 100% of expected time). Given this report's focus on full-time enrollees in two- and four-year programs, findings are related to students completing at 100% time, unless otherwise specified. Additional data for students completing at 150% and 200% of the expected time is available in the Dashboard, providing a broader context for understanding varied completion trajectories.

- Students who completed credentials had higher first-year GPAs, earned more credits in their first year, completed most of the courses they attempted in their first year, and enrolled more often in their second year than students who did not complete their credentials in a timely manner.
- Timely completers, on average, earned more than 90% of the credits they attempted in their first year and nearly all timely completers continued to be enrolled in a second year, either at the same or different institution.
- Second-year enrollment is particularly important for students pursuing an associate/certificate. The second-year enrollment rate for associate/certificate non-completers was 32 percentage points lower than that of students who completed. This gap is twice as large as the gap between bachelor's completers and non-completers (16 percentage points).
- The early success indicators of timely completers were higher than those of non-completers across all ethno-racial groups, genders, ages at enrollment, and fields of study. Furthermore, the disparities in early success indicators among demographic groups and fields of study were significantly smaller for completers. For example, the first-year credit completion ratio for associate/certificate completers ranged from 91% to 97% (6 percentage points) across the different ethnoracial groups, compared to a range of 56% to 76% (20 percentage points) for non-completers.
- The importance of early indicators in predicting completion varied by credential sought:
 - Associate/certificate seekers: Second-year retention (enrolling at the same institution) was the most important early success indicator followed by first-year credits earned, first-year GPA, and first-year credit completion ratio (CCR).
 - Bachelor's degree seekers: First-year GPA emerged as the most important of the early success indicators, followed by second-year retention, first-year CCR, and first-year credits earned.
- Most student demographic variables, such as race/ethnicity, gender, and age, are significant predictors of college completion, but their impact is less than that of early indicators. An exception is the first-year student major; for instance, majoring in social sciences has a stronger association with timely completion than credit completion ratio.

Section 1: Descriptive Analysis of Early Success Indicators by Completion Status

HOW DO EARLY SUCCESS INDICATORS DIFFER BY TIMELY COMPLETION STATUS?

Overall, timely completers had higher academic outcomes in their first year and enrolled at much higher rates for their second year than students who did not complete in a timely manner.

The averages for all early success indicators declined as students took longer to complete. However, timely completers, regardless of the length of time it took them to complete, earned higher average first-year GPAs, completed a greater percentage of their first-year courses, earned more first-year credits, and continued enrollment into their second year at a higher rate than those who did not complete (see Table 1).

Given this report's focus on full-time enrollees in two- and four-year programs, findings are related to students completing at 100% time, unless otherwise specified. Additional data for students completing at 150% and 200% of the expected time is available in the Dashboard, providing a broader context for understanding varied completion trajectories.

Table 1. Average Early Success Indicators by Completion at 100% Time and Degree Sought

Early Success Indicator	Associate/Certificate Seekers		Bachelor's Seekers	
	Completers	Non-Completers	Completers	Non-Completers
Average First-Year CCR	95.1%	65.5%	96.1%	78.4%
Average Credits Earned in First Year	30.0	16.7	29.8	22.4
Average First-Year GPA	3.34	2.27	3.45	2.61
Second-Year Enrollment Rate	100.0%	67.6%	99.9%	84.1%
Retention Rate	97.9%	60.8%	94.9%	68.9%
Transfer Rate	2.1%	6.9%	5.1%	15.2%
Student Count	16,107	87,086	85,346	109,071

Timely bachelor's completers generally had slightly higher early success indicators than timely associate/certificate completers.

Both bachelor's and associate/certificate completers had higher early indicator averages than non-completers. Additionally, bachelor's completers had higher early success indicators than associate/certificate completers, although the gaps were much smaller in comparison to the gaps between completers and non-completers. For example, the difference between bachelor's completers and associate/certificate completers' first-year CCR was about 1 percentage point, compared to double digit percentage point differences between completers and non-completers (Bachelor's seeking: 17.7 percentage points; Associate/certificate seeking: 29.6 percentage points).

Completers and non-completers had the largest gaps in first-year credit completion ratio and second-year enrollment. Timely completers successfully completed more than 95% of their first-year courses and nearly all enrolled in a second year.

More than 70% of timely completers had a first-year CCR of 100%, meaning they successfully completed all the courses they took (Bachelor's seekers: 75.8%; Associate/certificate seekers: 72.8%). Additionally,

nearly all completers enrolled in their second year (Bachelor's seekers: 99.9%; Associate/certificate seekers: 100.0%). In contrast, less than 40% of non-completers had a first-year CCR of 100% (Bachelor's seekers: 38.3%; Associate/certificate seekers: 27.0%) and less than 85% enrolled for a second year. Second-year enrollment rates for non-completers also varied more by credential sought (Bachelor's seekers: 84.1%; Associate/certificate seekers: 67.6%; see Table 1).

The disparity in second-year enrollment rates between those who completed their programs on time and those who did not was more pronounced for associate/certificate seekers (32 percentage points) compared to bachelor's degree seekers (16 percentage points).

Second-year enrollment rates were very high for completers across all time frames, with only 1% of completers, at the most, not enrolling again (see Dashboard: Early Success Indicators by Credential Completion). Bachelor's completers had a second-year enrollment rate of 99.9 percent, compared to 84.1 percent for non-completers. For associate/certificate seekers, the difference was 32 percentage points (100% vs. 68%). This indicates second-year enrollment may play even more of a role in timely completion for associate/certificate seekers than bachelor's seekers.

Students who completed in a timely manner had higher early indicators than non-completers, regardless of race, gender, age at entry, or major field of study. Early indicator gaps across demographic groups and fields of study for timely completers were narrower compared to non-completers.

While completion rates varied by race, age at entry, gender, and major field, the early success indicators of timely completers were consistently higher than non-completers, regardless of ethno-racial group, gender, age at enrollment, or field of study. Additionally, the differences in early success indicators among demographic groups and fields of study were notably smaller for completers compared to non-completers (see an example in Table 2 and all comparisons in Dashboard: Early Success Indicators by Credential Completion).

Table 2. Range of Average Credit Completion Ratios by Student Demographic and Major

	Associate/Certificate Seekers		Bachelor's Seekers	
	Completers	Non-Completers	Completers	Non-Completers
Race/Ethnicity	91.1%-97.1% (6.0 pp)	56.1%-76.1% (20.0 pp)	94.6%-97.8% (3.2 pp)	72.0%-82.8% (10.8 pp)
Gender	94.6%-95.6% (1.0 pp)	63.9%-67.7% (3.8 pp)	96.1%-96.3% (0.2 pp)	77.3%-80.0% (2.7 pp)
Age at Entry	95.1%-95.7% (0.6 pp)	60.8%-66.3% (5.5 pp)	94.2%-98.0% (3.8 pp)	63.2%-78.8% (15.6 pp)
Major	94.7%-96.1% (1.4 pp)	61.6%-71.4% (9.8 pp)	95.0%-96.8% (1.8 pp)	75.7%-81.8% (6.1 pp)

Section 2: Mixed Effects Logistic Regression Analysis

Above we have discussed descriptive differences in early success indicators between timely completers and non-completers. However, without further statistical analysis it is unclear how these factors function together and if they outweigh the impact of student characteristics and institutional differences on students' probability of timely completion.

For this reason, we also analyzed the data by estimating a mixed-effect logistic regression, a statistical method used to predict the probability of an event (timely completion) while accounting for both fixed effects (predictors like early success indicators and demographics variables) and random effects (variability between institutions). Six regression models were estimated based on the credential sought (associate/certificate and bachelor's) and the completion timeframe (100%, 150%, and 200%).

The mixed-effect model was estimated with students nested within institutions and includes two groups of independent variables: all early success indicators (first-year GPA, first-year CCR, first-year credits earned, and second-year retention) and a set of student demographic variables (gender, age at entry, and race/ethnicity) and the student's first-year major. Note that second-year retention (enrolling for a second year at the same institution) is a subset of second-year enrollment (enrolling for a second year regardless of the institution), which was assessed above in the descriptive statistics. For information on why this early indicator was adjusted slightly for the regression model, model fit, variable selection, and institutional level differences, see Methodological Notes: Regression Model.

Multiple statistical values can be used to interpret the impact of a predictor on timely completion and have been provided on the Dashboard and in the Methodological Notes. For ease of interpretation, we will focus on the predicted probability of a variable. The predicted probability tells us, holding everything else constant, what the increased or decreased probability of timely completion would be when that variable is present (categorical variables) or increases by one standard deviation (continuous variables; see Methodological Notes for more information on model interpretation).

HOW WELL DO THE EARLY SUCCESS INDICATORS PREDICT TIMELY COMPLETION FOR FALL FIRST-TIME, FULL-TIME STARTERS?

All four early success indicators tested in the models were significant predictors of timely completion. Additionally, early success indicators were stronger predictors of timely completion than student demographic variables, regardless of credential sought and completion timeframe.

All four early success indicators are statistically significant predictors of timely completion with predicted probability values from 57% to 91%, depending on the indicator and model (see Table 3). For example, the probability of a student completing an associate/certificate was 91% higher when the student continued at the same institution into their second year (i.e., were retained), compared to students who were not retained, holding all other early indicators and student characteristics the same. Although most of the demographic variables (i.e., race, age at entry, and gender) were found to be statistically significant predictors of timely completion, the strength of their predictive ability was much smaller with predicted probability values from 32% to 56%.

The impact of early success indicators in predicting timely completion varied by credential sought. For associate/certificate seekers enrolling at the same institution for a second year was the most important early success indicator, followed by first-year credits earned, first-year GPA, and first-year credit completion ratio (CCR).

For associate/certificate seeking students, the impact of early success indicators had greater variation; the probability of timely completion was 57% to 91% greater depending on the indicator. In contrast, early success indicators for bachelor's seeking students increased the probability of timely completion by a range of 66% to 79%.

Second-year retention was an important predictor for both groups; however, it was even more pronounced for associate/certificate seekers, similar to what was seen in descriptive statistics for second-year enrollment at any institution. When students were retained at the same institution, they were 91% more likely to complete than students who transferred or did not re-enroll for a second year. This finding should be interpreted with caution, though, as this indicator evaluates second-year enrollment in a program that is expected to take two years or less. Regardless, retention remained the strongest predictor of associate/certificate completion across all timeframes (see Table 3).

The number of credits earned was the next most important early success indicator for associate/certificate seekers; the probability of completing was 81% higher for students who earned an additional 10 credits (one standard deviation in credits earned). These findings suggest that interventions for associate/certificate seeking students will likely be most fruitful if focused on supporting second-year student retention and earning more credits in their first year. The average number of credits earned for associate/certificate completers was 30 credits and over 80% of completers earned at least 25 credits in their first year.

Table 3. Regression Results for Completion at 100% Time

		Associate/Certificate Completion		Bachelor's Completion	
		Predicted Probability	p-value	Predicted Probability	p-value
Early Indicators	CCR	57%	0.000	67%	0.000
	Credits Earned	81%	0.000	66%	0.000
	GPA	70%	0.000	79%	0.000
	Retention	91%	0.000	76%	0.000
First-Year Major	Business	56%	0.000	69%	0.000
	Education	49%	0.476	54%	0.000
	Social Sciences	61%	0.000	70%	0.000
	STEM	38%	0.000	50%	0.508
	Other/Unknown	57%	0.000	59%	0.000
Race/Ethnicity	White	55%	0.000	56%	0.000
	Black	49%	0.376	47%	0.000
	Asian	47%	0.007	53%	0.000
	Nonresident Alien	45%	0.004	54%	0.000
	Other	50%	0.890	49%	0.173
	Unknown	52%	0.270	53%	0.021
Age	20-24	47%	0.002	45%	0.000
	Older than 24	48%	0.108	32%	0.000
Gender	Male	49%	0.139	40%	0.000
	Unknown	55%	0.000	45%	0.000
Intercept		1%	0.000	6%	0.000

Note. Bolded values are statistically significant ($p < 0.05$). The reference group is students with average CCR, credits earned, and GPA, who were not retained, had a Liberal Arts major, and are Hispanic, a woman, and who started at their institution when they were younger than 20.

GPA emerged as the most important predictor for bachelor’s degree seekers followed by second-year retention, first-year CCR, and first-year credits earned.

In contrast to associate/certificate seekers, first-year GPA was the strongest predictor of completion for bachelor’s degree seekers. A GPA increase of 1.05 (one standard deviation) is associated with a 79% higher probability of completing a bachelor’s in four years. Retention followed closely behind—the probability of completing was 76% higher for students who enrolled for a second year at the same institution, compared to students who either transferred or did not enroll in their second year.

While all early success indicators are important, the bachelor’s seeking student model implies interventions should focus on increasing students’ first-year GPA and retention into their second year. Over 80% of bachelor’s seeking completers earned at least a 3.0 with an average first-year GPA of 3.5.

The first-year major a student chooses to pursue can be as predictive of timely completion as the early success indicators. Social Science and Business majors had the strongest association with timely completion of the major fields analyzed.

Major fields, with two exceptions (STEM for associate/certificate seekers and Education for bachelor’s seekers), were statistically significant predictors across all models. In some cases, major field impacted the probability of completion at the same, or higher magnitude of early indicators, which was not the case for the other student-related variables (demographics), as mentioned above (see Table 3).

Compared to the reference group of Liberal Arts majors, Social Sciences majors had a probability of earning an associate/certificate in two years that was 61% higher. Additionally, being a Social Sciences major was the 5th strongest predictor of earning an associate/certificate in two years over the early success indicator of CCR, the 7th strongest predictor.

For bachelor’s seekers, the probability of earning a bachelor’s in four years was 70% higher for Social Science majors, compared to Liberal Arts majors. Being a Business major also had a similar association—the probability of earning a bachelor’s in four years was 69% higher for Business majors, compared to Liberal Arts majors. Majoring in Social Sciences or Business (compared to Liberal Arts) were the 3rd and 4th strongest predictors, respectively, surpassing CCR and credits earned at the 6th and 7th spots, respectively (see Table 4).

Table 4. Rank for Top 10 Strongest Predictors of Completion at 100% Time

Rank	Associate/Certificate Completion	Bachelor’s Completion
First	Retention	GPA
Second	Credits Earned	Retention
Third	GPA	Major: Social Sciences
Fourth	Major: STEM*	Major: Business
Fifth	Major: Social Sciences	Age: Older than 24*
Sixth	Major: Other/Unknown	Credit Completion Ratio
Seventh	Credit Completion Ratio	Credits Earned
Eighth	Major: Business	Gender: Male*
Ninth	Race: Nonresident Alien*	Major: Other/Unknown
Tenth	Race: White	Race: White

Note. Early success indicators are bolded. *If a student had a lower probability of completing based on a characteristic (e.g., having a STEM major), that is noted with an asterisk.

However, this advantage diminishes over extended graduation timeframes. By the 200% timeframe, Social Sciences and Business majors were only 57% and 61% more likely to earn a bachelor's degree than Liberal Arts majors.

Section 3: Special Analysis of Bachelor's Earning Associate/Certificate Seekers

Nearly 10,000 students who were seeking an associate/certificate upon entry completed a bachelor's degree within eight years but did not earn an associate/certificate credential. This section further analyzes the early success indicators for these bachelor's completers compared to the completers who earned the degree they originally pursued.

Table 5. Average Early Success Indicators for Bachelor's Earning Associate/Certificate Seekers

	Bachelor's Earning Associate/Certificate Seekers	Associate/Certificate Completers (200% Time)	Bachelor's Completers (200% Time)
Average First-Year CCR	90.2%	91.1%	94.0%
Average Credits Earned in First Year	25.3	26.8	28.7
Average First-Year GPA	3.18	3.20	3.32
Second-Year Enrollment Rate	98.3%	99.1%	99.6%
Retention Rate	68.5%	94.9%	92.1%
Transfer Rate	29.8%	4.2%	7.5%
Student Count	9,947	36,250	133,858

Early success indicators for bachelor's earning associate/certificate seekers were comparable to those of associate/certificate completers at 200% time.

Bachelor's earning associate/certificate seekers had higher early indicators than non-completers, regardless of credential sought, but their indicators were still lower than those of completers across all time groups. The early success indicators for these students, however, were most similar to those of associate/certificate seekers who completed in 4 years or less (200% time; see Table 5).

Nearly thirty percent of associate/certificate seekers who earned a bachelor's transferred to another institution by their second year.

Like second-year enrollment rates for other completers, almost all bachelor's earning associate/certificate seekers (98.3%) continued enrollment into their second year. However, these students also had the highest transfer rates of all student groups; almost thirty percent (29.8%) enrolled at a different institution in their second year (see Table 5).

Section 4: Summary

This report underscores the critical importance of early success indicators in timely credential completion among full-time post-secondary students. Key indicators, including first-year grade point average (GPA), credit completion ratio (CCR), credits earned, and second-year enrollment, most

specifically retention at the same institution, emerged as significant predictors, with their influence consistently exceeding that of demographic variables.

The relative importance of these indicators varied by credential type; retention was particularly critical for associate and certificate seekers, while first-year GPA was the strongest predictor for bachelor's degree seekers. Additionally, students' major field demonstrated a substantial impact on timely completion, often rivaling or surpassing the influence of other predictors.

These findings emphasize the need for targeted, evidence-based interventions that prioritize early academic achievement, support retention, and address program-specific challenges to improve completion outcomes.

Methodological Notes

Cohort Composition

Data for this report comes from the National Student Clearinghouse Postsecondary Data Partnership (PDP). Institutions actively opt-in and pay a small annual fee to participate in PDP. Therefore, while this report can uniquely characterize the early success indicators and completion outcomes for a large sample of students, this is not a nationally representative set of institutions. Participating institutions may vary from others in their student outcomes, demographics, programs, and services. No findings in this report should be considered representative of the national population of students or of institutions. Data was reported as of September 26, 2024. See Table M1 for student and institution counts by sector.

Table M1. Student and Institution Counts by Sector

Sector	Institution		Student	
	Count	Share	Count	Share
Public 4-year	99	33.9%	187,608	61.0%
Private 4-year	30	10.3%	13,221	4.3%
PAB*	14	4.8%	10,396	3.4%
Public 2-year	149	51.0%	96,323	31.3%
Total	292		307,548	

*PAB stands for Primarily Associate Degree Granting Baccalaureate Institutions. These are institutions that are considered 4-year institutions but primarily grant associate degrees (see Institution Sector Classification for more information).

The cohort consists of first-time in college, credential-seeking (associate/certificate or bachelor's degree), full-time students who first enrolled during the fall semester of the 2016-17 academic year. To be included in the analysis, each student must have complete information reported on completion status and all early success indicators.

Additionally, to account for potential reporting errors, students who had any early success indicators outside of the expected ranges, per PDP data audit expectations and guidebook definitions, were excluded from the sample.

Students whose indicators fell within these ranges were included:

- First -Year Credit Completion Ratio: 0-100%
- First-Year Credits Attempted: 0-99 credits
- First-Year Credits Earned: 0-99 credits
- First-Year GPA: 0-4.0
- Second-Year Enrollment: Had to be identified in one of these categories:
 - Enrolled for a second year at the same institution or completed a credential by the second year
 - Enrolled for a second year at another institution or completed a credential by the second year
 - Did not enroll for a second year and did not complete a credential by the second year

Sample Representation

The final sample consists of 307,566 unique students. Table M2 captures institutional and student counts.

Table M2. Student and Institution Counts by Credential Sought

Student Group	Student Count	Institution Count
Associate/certificate Seeking	113,135	191
Bachelor's Seeking	194,413	141

Note: The institution counts when totaled equals more than 292 as some institutions offer both associate/certificate and bachelor's degrees (40 institutions in this sample).

The sample presented in this report mirrors the gender and age distribution of the broader student population enrolled during the fall semester of the 2016-17 academic year (refer to Dashboard: Sample Demographics). However, as this sample constitutes less than two percent of the total undergraduate enrollment for Fall 2016, caution is warranted when extrapolating these findings to a national context (see [Fall 2016 CTEE Report](#)).

First-Year Credential Sought

A student's credential-seeking status is based on their "Degree Type Sought" value in an institution's PDP submission file. If degree type sought is not populated in the Course Extended file, or if this field was populated in historical data from a secondary data source that does not have values that can be mapped with the values in the Course Extended File submission guide, then this is populated from NSC enrollment records.

The PDP dataset defines completion as bachelor's completion or completing a certificate or associate degree; certificate and associate degree completion are not tracked separately. For this reason, the following credential-seeking categories were combined to create two main categories:

- Associate/certificate degree seeking students who were comprised of students with the following first-year credential sought categories:
 - Less than one-year certificate, less than Associate degree
 - One-to-two-year certificate, less than Associate degree
 - Two-to-four-year certificate, less than Bachelor's degree
 - Associate Degree
- Bachelor's degree seeking students

These credential type options are excluded from the sample as non-credential seeking categories:

- Undergraduate Certificate or Diploma Program
- Post Baccalaureate Certificate
- Master's Degree
- Doctoral Degree
- First Professional Degree
- Graduate/Professional Certificate
- Non-Credential Program (Preparatory Coursework/Teach Certification)

Early Success Indicators

First-Year Credits Earned

Course information for all students in the cohort is provided by the reporting institution and includes data such as the course name, course type (i.e., college-level, college developmental, adult basic education), course grade and credit hours attempted.

The total credit hours earned is the sum of credit hours earned for each course during the first year. The course grade and credit hours attempted for each course are utilized to calculate the credit hours earned. Failing grades and course withdrawal mean a student did not receive credit for a course, yielding zero credit hours.

Earned credit hours are calculated for each course regardless of course type, meaning it is possible that courses that do not count toward credential completion are being included (i.e., developmental college courses not applicable towards credential completion). For this reason, it is likely that the results for indicators used in this report may be slightly higher than they would be if limited to only certain course types. First-year credits earned is a continuous variable ranging from 0 to 99 as the audit check for PDP flags students earning 100 or more credits a year.

First-Year Credit Completion Ratio

The credit completion ratio (CCR) is the ratio of credits earned to credits attempted in a student's first year of enrollment. That is, among all credits attempted in their first year, the CCR tells what proportion of those credits the students earn (i.e., receive credit for). The overall CCR is calculated as:

$$\frac{\textit{Credit Hours Earned in First Year}}{\textit{Credit Hours Attempted in First Year}}$$

CCR can provide insights into students' overall course completion and measure the efficiency of students' movement through coursework. First-year CCR is a continuous variable ranging from 0 to 100%.

First-Year Grade Point Average

The first-year grade point average (GPA) is a student's GPA at the end of the first academic year based on the course data provided by the institutions. First-year GPA is a continuous number with values between 0 and 4.0.

Second-Year Enrollment

A student is considered to have second year enrollment if they completed a credential by the second year of enrollment or have continued enrollment into their second academic year. This categorical variable consists of three outcomes:

- **Retained:** Continued into their second year at the same institution or earned a credential
- **Transferred:** Continued into their second year at another institution or earned a credential
- **Not Enrolled:** Did not continue into their second year and did not earn a credential

As PDP tracks retention and transferring over the year and not term to term, it is possible for students to meet both outcomes in a year. When this happens, the PDP dashboards prioritize retention. Following suit, 628 students who were identified as having both transferred and been retained were categorized as retained.

Completion Metrics

Completion in this report is defined by the student having earned the credential they initially sought, at any institution, within a specific time frame. Specifically, a student initially seeking a bachelor's would be considered to have completed their credential only if they earned a bachelor's in a specific time frame, while a student seeking an associate/certificate would be considered to have completed only if they earned an associate/certificate degree in a specific timeframe. Students who did not earn a credential at all or earned the credential sought outside of the specified time frame are defined as not completing.

Time to completion is calculated as follows and rounded up to the nearest whole number:

$$\frac{\text{First Credential Award Date} - \text{Cohort Begin Date}}{365}$$

Table M3. Time Frame for Completion

Student Group	Time to Completion		
	100%	150%	200%
Associate/certificate Seeking	2 years or less	3 years or less	4 years or less
Bachelor's Seeking	4 years or less	6 years or less	8 years or less

Bachelor's Earning Associate/Certificate Seekers

In the case of associate/certificate seekers, a total of 9,947 students earned only a bachelor's degree at some point within the eight years tracked without any associate/certificate completion. Based on the definition of completion for this report, these students were not identified as completing within the time frame. Rather, they were identified as a separate group, Bachelor's earning associate/certificate seekers. These students are analyzed in a special analysis in this report.

Student Demographics

Race

Student race is reported to the PDP by institutions when students first enroll. The categories include White, Hispanic, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Two or More Races, Black or African American, Nonresident Alien, Asian, and Unknown. In this report, 1.8% of the students' race was reported as unknown.

Gender

Student gender can be reported to the PDP by institutions when students first enroll. The categories are Male, Female, Prefer not to specify, Nonbinary, intersex, and gender-nonconforming, and Unknown. Prefer not to say and Unknown were combined to be Unknown and none of the students in this report were identified as Nonbinary, intersex, and gender-nonconforming, resulting in the following three categories presented in this report:

- Male
- Female
- Unknown

If gender is not reported to PDP, it is imported from data reported to the Clearinghouse’s enrollment reporting service. If gender is not reported to the PDP or enrollment reporting, it is imputed.

The Research Center developed an imputation process based on first names. Previously submitted name gender pairs throughout the Clearinghouse database are used to determine the probability of any first name being associated with either gender. To improve the accuracy of the imputation process, the Research Center also draws on name-gender data from the Social Security Administration (SSA) and the U.S. Census Bureau. Because the Clearinghouse collects transactional data, its data contain many more unique first names than other sources. The imputation used only those pairs in which the name occurred in at least two instances and was associated with a single gender in at least 95 percent of the instances. The SSA and Census data sets were used to ensure that name-gender pairs were consistent across every data set in which they occurred and to enhance the imputation process by contributing name-gender pairs that did not occur in the Clearinghouse data.

If gender was reported as Unknown to PDP, not reported to PDP in enrollment reporting, and could not be imputed, the value was left Unknown. In this report, 6.6% of the students’ gender was reported as Unknown.

Age at Entry

A student’s age at entry is calculated as:

$$\frac{\text{Cohort Begin Date} - \text{Date of Birth}}{365.25}$$

Once rounded to the nearest whole number, ages are grouped into 3 categories:

- 20 and younger
- 21-24
- Older than 24

First-Year Major of Study

A student’s first-year major is based on the program of study CIP code reported by the cohort institution in their first academic year. The first-year program of study was missing for 22,556 students. In 44% of these cases (9,945), the student’s first semester program of study was available and was imputed as their first-year major.

For subgroup analyses by major field of study, PDP institutions report the six-digit CIP code for each student’s program of study. These CIP codes were collapsed by the leading two (subject) digits and organized by the U.S. Department of Education’s CIP taxonomy. Majors classified in “STEM” include CIPs: 14- Engineering, 26- Biological and Biomedical Sciences, 27- Mathematics and Statistics, and 40- Physical Sciences. These follow the U.S. Department of Homeland Security’s classification of STEM fields.

Institution Sector Classification

This report defines institution sectors based on the classification at the time the cohort started postsecondary education (2016-2017 academic year). Sector classifications are primarily based on the 2015 [Carnegie Basic Classification](#). Carnegie Classification and IPEDS sector designations align for the most part, but when there are differences, we follow the Carnegie Classification. These discrepancies mostly impact Primarily Associate Degree Granting Baccalaureate Institutions (PABs; see below). When a Carnegie Classification is missing for an institution, we utilize the institution's IPEDS sector (2016-2017 academic year) where available, and, finally, the sector reported by the institution to the Clearinghouse.

Primarily Associate Degree Granting Baccalaureate Institutions (PABs)

As more and more institutions that previously focused solely on granting associate degrees have begun to offer bachelor's degree programs, there has been a surge in IPEDS reclassification of 2-year institutions as 4-year institutions, since IPEDS assigns 2- or 4-year designations based on program offerings. However, many of these reclassified institutions still confer most awards at the associate degree level. These are considered primarily associate degree granting baccalaureate (PAB) institutions. We utilize the 2015 Carnegie Basic Classification to identify PABs. PABs are defined as institutions that offer at least one baccalaureate degree program and award more than half of their degrees at the associate level. These institutions are made up of two subcategories:

- Baccalaureate/Associate Colleges - Associate Dominant (code 14): institutions that award 90 percent or more of degrees at the associate level, or
- Baccalaureate/Associate Colleges - Mixed Baccalaureate/Associate (code 23): institutions that award more than 50 but less than 90 percent of degrees at the associate level.

Regression Model Creation

Data Filtering and Manipulation

The dataset used for descriptive statistics was also utilized for the model. To ensure alignment with the study's focus on timely completion, Bachelor's Earning Associate/Certificate seekers were excluded from the analysis, as they did not fit within the scope of the timely completion timelines (see Completion Metrics above for more information).

Please review the table below for the sample sizes entering the BA and UCAA models after excluding these students.

Table M4. Sample Sizes After Exclusion of Bachelor's Earning Associate/Certificate Seekers

Student Group	Number of Students	Number of Unique Schools
Associate/certificate Seeking	103,196	187
Bachelor's Seeking	199,423	137

Handling Outliers

As mentioned in the Cohort Composition section of the Methodological Notes, the continuous variables (CCR, GPA, and Credits Earned) were restricted to predetermined ranges, effectively removing outliers from the dataset and ensuring that only valid, relevant data was included in the analysis.

Missingness

The subset of data used for model training contained minimal missingness. For demographic variables that included unknown categories, students were *not* excluded from the analysis. Instead, we created specific “unknown” categories for most of these demographic variables. Please review table below for the number of observations entering the UCAA and BA models.

Variable Recoding

- **Categorical Variables**

Major: The major variable was recoded into six dummy variables based on the first two digits of the six-digit CIP code associated with each student’s record. These categories include Liberal Arts (reference group), STEM (see First-Year Major of Study above for more information), Social Sciences, Business, Education, and Other. Unlike other dummy-coded variables, in which missing data are represented by a distinct category, the Other major dummy includes the 12,412 students with missing major information. We initially tested model iterations where students with missing major information were assigned a separate dummy variable; however, this approach resulted in high collinearity with other major categories, leading us to group missing and other cases together.

Age: The age variable was recoded into three dummy variables: students 20 and younger (reference group), students 20-24, and those older than 24. Students 20 and younger were selected as the reference group because they represent the largest proportion of the sample.

Race: The race variable was recoded into seven groups: Hispanic (reference group), Black, White, Asian, Unknown, Non-Resident, and Other. The “Other” dummy variable includes students who identified as Two or More Races, Native Hawaiian or Other Pacific Islander, or American Indian or Alaskan Native, which accounts for 5.3% of the dataset.

Gender: The gender variable was recoded into three dummy variables: female (reference group), male, and missing. The gender missing dummy variable encompasses all students with missing gender information (20, 285 students) and those who selected “prefer not to say” (15 students)

Retention: As outlined in the Cohort Composition section of the Methodological Notes, second-year enrollment was initially divided into three categories: transfer, no-enrollment, and retention. We originally included each category in the model as a separate dummy variable, but high collinearity was observed between the transfer and retention variables. To address this, we combined the transfer and no-enrollment groups into a single reference dummy variable and included retention as the primary comparison group in the model. It is important to highlight that retention operates differently within the UCAA and BA models due to the varying timeframes associated with each credential type. For Bachelor’s students, retention is defined as re-enrollment in the second year, representing 25% of the credential’s typical completion timeframe. In contrast, for Associate/Certificate-seeking students, re-enrollment in the second year constitutes 50% of the anticipated completion timeframe. This difference in the ratio of retention timeframe to completion timeframe results in retention being a notably stronger predictor of completion for UCAA students than for Bachelor’s students.

Generally, we used the largest sub-groups as the reference group. However, although the "Other" major sub-group is larger than the Liberal Arts sub-group, we chose Liberal Arts as the reference group to provide clearer contextual interpretation. Additionally, we used Hispanic students, the largest group after White students, as the reference category.

- **Continuous Variables**

To ensure that the continuous variables (first-year GPA, first-year CCR, and first-year Credits Earned) were comparable in range and could be effectively used in the model, these variables were standardized. Each continuous variable was scaled and centered to have a mean of 0 and a standard deviation of 1. This transformation was performed to facilitate model convergence and improve the overall interpretability of the results. Please see table below for the values corresponding to a single standard deviation increase for each of the continuous variables included.

Table M5. Changes Corresponding to One Standard Deviation Increase

Continuous Variable	Standard Deviation
First-Year Credits Earned	10.39 credits
First-Year GPA	1.05 points
First-Year Credit Completion Ratio	28 percentage points

Model Design

Six mixed-effects logistic regression models were conducted, representing different degree sought and completion cutoffs. Specifically, three models were developed for Bachelor's degree seekers at three distinct timeframes, and three additional models were created for Associate/Certificate degree seekers. The following equation illustrates the model structure, where students (i) are nested within schools (j).

$$\begin{aligned}
 \text{Logit}(\text{Timely Completion} = 1) &= \beta_0 + \beta_1 \text{GPA}_{ij} + \beta_2 \text{CCR}_{ij} + \beta_3 \text{Credits}_{ij} + \beta_{4-8} \text{Major}_{ij} + \beta_{9-10} \text{Gender}_{ij} \\
 &+ \beta_{11-12} \text{Age}_{ij} + \beta_{13-18} \text{Race}_{ij} + \beta_{19} \text{Retention}_{ij} + \alpha_j + \epsilon_{ij} \\
 &\sim N(0, \sigma_\epsilon^2)
 \end{aligned}$$

Model Testing

- **Multicollinearity: VIF testing**

A Variance Inflation Factor (VIF) test was conducted to assess the presence of collinearity among the fixed-effects variables in the model. This analysis informed the selection of the final model variables. Initially, we included variables for persistence (students enrolled at the same institution in their second year), retention (students enrolled in any institution in the second year), and unknown major. However, these variables exhibited high collinearity (VIF > 10) and were subsequently removed from the final model. The VIF values for the final model were all less than 3.5.

- **Model Fit:**

Conditional and Marginal R^2 : Marginal R^2 indicates the proportion of variance explained by the fixed effects, reflecting the contribution of the model's predictors. On the other hand, conditional R^2 accounts for the variance explained by both the fixed and random effects, offering a more comprehensive measure of model fit (Nakagawa & Schielzeth, 2013). These metrics are particularly useful in mixed-effects models, as they help to assess the contribution of both fixed and random components to the variability in the data. Generally, marginal R^2 values between 0.2 and 0.4 suggest a moderate explanatory power of the fixed effects, while values above 0.4 might indicate stronger explanatory power. However, very high R^2 values should be interpreted with caution, as they might also indicate the model is overfitting the data. These ranges of R^2 values provide a nuanced perspective on model fit and the relative contributions of fixed and random components, but careful consideration is needed to avoid overinterpretation. As seen in the table below, the models appear to have strong explanatory power.

Table M7. Marginal R^2 and Conditional R^2 Values

Credential Sought	Model of Timely Completion	R^2 Values	
		Marginal	Conditional
Associate/ Certificate	Completion within 100% time (2 years)	0.71	0.78
	Completion within 150% time (3 years)	0.67	0.72
	Completion within 200% time (4 years)	0.63	0.68
Bachelor's	Completion within 100% time (4 years)	0.58	0.67
	Completion within 150% time (6 years)	0.52	0.60
	Completion within 200% time (8 years)	0.50	0.58

Due to the high Marginal and Conditional R^2 values observed, particularly in the UCAA 100% model, there was concern regarding potential overfitting. To address this, we developed alternative models by applying feature selection methods and excluding variables with high predictive power, such as retention. Although these alternative models yielded slightly lower Marginal and Conditional R^2 values and better model-fit metrics, the differences were minimal. We ultimately opted to retain the more complex models to better capture the nuanced relationships among variables.

Model Fit Metrics: We conducted multiple iterations of the models, altering input variables to optimize model fit based on AIC, BIC, Log Likelihood, and Deviance values. While the final model did not yield the best AIC or BIC, the differences in these criteria between the final model and alternative specifications were marginal. Thus, we selected a more complex model to preserve interpretive insights into the relationships among the variables. Notably, the removal of the retention variable led to the largest reduction in model variation, underscoring its significant contribution to explaining variability across models.

Model Interpretation

To ensure clarity in interpreting the results of our mixed-effects logistic regression model, we provide the following definitions of key concepts:

1. **P-value:** The p-value is a statistical measure used to assess the evidence against the null hypothesis. It represents the probability of observing the data, or more extreme outcomes, under the assumption that the null hypothesis is true. A p-value of 0.05 or lower is typically

considered statistically significant, indicating that the observed effect is unlikely to have occurred by random chance.

2. **Log Odds Coefficient:** In logistic regression, the relationship between the dependent variable and the predictors is modeled in terms of log odds. The odds of an event occurring are defined as the ratio of the probability of the event to the probability of it not occurring. The log odds are the natural logarithm of the odds.

$$\text{Odds} = \frac{P}{1-P}$$

$$\text{Log Odds} = \ln(\text{Odds})$$

Negative log odds correspond to probabilities less than 50%, whereas positive log odds correspond to probabilities greater than 50%. For example, retention has a log odds coefficient of 2.33 for associate/certificate completion in two years or less. This means a student who is retained for their second year, keeping every other variable constant, has a higher probability of completing an associate/certificate in two years or less than those who are not retained.

3. **Odds Ratio:** The odds ratio (OR) is a measure used to quantify the association between a predictor variable and the odds of an event occurring. It is the ratio of the odds of the event in one group relative to another. An odds ratio of 1 indicates no effect, meaning the odds of the event are the same across groups. An odds ratio greater than 1 suggests an increased probability of the event in the exposed group, while an odds ratio less than 1 indicates a decreased probability compared to the reference group. For example, retention has an odds ratio of 10.25 for associate/certificate completion in two years or less. This means a student who is retained for their second year, keeping every other variable constant, has an odds of completing an associate/certificate in two years or less that is 10.25 times that of those who are not retained.

$$OR = \frac{\text{Odds of event in Group 1}}{\text{Odds of event in Group 2}} = \frac{\frac{P_1}{1-P_1}}{\frac{P_2}{1-P_2}}$$

4. **Predicted Probability:** Predicted probability is derived from the odds ratio using the following equation:

$$\text{Probability} = \frac{OR}{1 + OR}$$

Converting odds ratios to predicted probabilities allows for a clearer and more interpretable presentation of results. The predicted probability tells us how much a variable changes the probability of the expected outcome, in this case timely completion. Variables with a predicted probability greater than 50% indicate a higher probability of completion, whereas those with a predicted probability less than 50% indicate a lower probability of completion. For example, retention has a predicted probability value of 91% for associate/certificate completion in two years or less. This means a student who is retained for their second year, keeping every other variable constant, has a probability of completing an associate/certificate in two years or less that is 91% higher than those who are not retained.

References

Al-Haddad, S., Boone, R., & Campbell, E. (2018). Understanding graduation rates at higher education institutions: A forecasting model. *International Journal for Business Education*, 158, 10-23. <https://www.ijbe.online/2018-understanding-graduation-rates-at-higher-education-institutions.html>

Bound, J., Lovenheim, M. F., & Turner, S. (2012). Increasing time to baccalaureate degree in the United States. *Education Finance and Policy*, 7(4), 375-424. https://doi.org/10.1162/EDFP_a_00074

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: Analysis and applications*. Cambridge University Press.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133–142. <https://doi.org/10.1111/j.2041-210X.2012.00261.x>

Ortagus, J. C., Skinner, B. T., & Tanner, M. J. (2021). Investigating why academically successful community college students leave college without a degree. *AERA Open*, 7(1), 1-17. <https://doi.org/10.1177/23328584211065724>

Suggested Citation

Gardner, A., Ibrahim, M., Karamarkovich, S., Holsapple, M., and Shapiro, D. (December 2024), *PDP Insights: Early Indicators of Timely Completers*, Herndon, VA: National Student Clearinghouse Research Center.

The authors thank National Student Clearinghouse staff members Benjamin Berg, Jennifer Causey, Hee Sun Kim, Benjamin England, and Marty Schmidt for their assistance with this publication.