

Methodological Notes

This report examines six-year and eight-year college student success outcomes, focusing primarily on degree and certificate completion of a cohort of first-time-in-college, degree-seeking students who started their postsecondary education at U.S. colleges and universities in the fall of 2014 and fall of 2012 through June 30, 2020. Outcomes examined include completion (i.e., receipt of any postsecondary credential by the end of the study period), major at completion, persistence (i.e., having enrollment records at any postsecondary institution during the last year of the study period), and stop-out without completion (i.e., having no enrollment records at any postsecondary institution during the last year of the study period). The report mainly focuses on students' first ever completions, with further distinctions drawn between completions awarded at the institution where a student first enrolled (his or her starting institution) and those awarded at an institution other than their starting institutions. For students who started at a two-year public institution, this report also presents an overview of their completions at a four-year institution, either as a first completion (i.e., those who completed at a four-year institution without having first earned a credential at a two-year institution¹) or as a subsequent degree after a first completion awarded in the two-year sector. In addition to results on degree and certificate completion rates by enrollment intensity, age group, gender, race and ethnicity, and starting institution type, the report includes results on completion across state lines and for students who started at multistate institutions.

¹ Throughout this report, "two-year institution" is used broadly to designate institutions offering both associate degrees and less-than two-year degrees and certificates.

Data Coverage

The National Student Clearinghouse currently collects data from more than 3,600 postsecondary institutions, which represent 97 percent of the nation's postsecondary enrollments in degree-granting institutions, as of 2019. The enrollment data used in this report provide an unduplicated headcount for the fall 2014 first-time college entering student cohort. Clearinghouse data track enrollments nationally and are not limited by institutional and state boundaries. Moreover, because this database is comprised of student-level data, researchers can use it to link concurrent as well as consecutive enrollments of individual students at multiple institutions.

Cohort Identification, Data Cut, and Definitions

This report examines completion, over a span of six years, for the cohort of first-time-in-college degree-seeking students who started their postsecondary studies at U.S. colleges and universities in the fall of 2014 for six-year outcomes as well as in the fall of 2012 for eight-year outcomes (through June 30th, 2020). To limit the cohort to first-time undergraduate students only, the study uses data from the Clearinghouse's enrollment reporting and DegreeVerify services to confirm that students included in the study fulfilled the following conditions:

1. Enrolled in a Title IV degree-granting institution in fall 2014, excluding territories outside the U.S. (e.g., Guam, Puerto Rico, or the U.S. Virgin Islands).
2. Did not have a previous enrollment record, as shown in the Clearinghouse data, prior to the first day of enrollment in the fall of 2014, unless the previous enrollment record was before the student turned 18 years old (qualified as former dual enrollment students);
3. Did not receive any degree or certificate from a postsecondary institution prior to the first day of enrollment in fall 2014, according to Clearinghouse data unless the award date was before the student turned 18 years old (dual enrollment);
4. Had at least one legitimate enrollment status throughout the study period; that is, enrolled for at least one term with full-time, part-time (i.e., halftime or less than half-time), or withdrawal status²;
5. Showed intent to seek a degree or certificate. That is:
 - For students who started at four-year institutions, enrolled at least one term with an intensity of half-time or higher.
 - For students who started at two-year institutions, either:

- Enrolled full time for at least one term before August 10, 2015, or
- Enrolled three-quarter time for at least one term or half time for any two terms before December 31, 2015³

²The Clearinghouse receives enrollment status data as full-time, half-time, less-than-half-time, withdrawal, or other statuses from its participating institutions.

³ We excluded 245,165 students who began at two-year institutions as non-degree-seeking students as a result.

Race and Ethnicity

The race and ethnicity categories include Asian, African American (of non-Hispanic origin), Hispanic, White, Other or two or more races. The Other category includes American Indian/Alaskan native, and Native Hawaiian or other Pacific Islander. Seventeen percent of the data are unknown or missing race/ethnicity.

Former Dual Enrollment Students

Beginning in the fall 2011 entering cohort, our definition of first-time student cohort excludes current dual enrollment students, impacting largely two-year institutions. Interpret the trends with caution for two-year starters because removing current dual enrollment students caused the completion rate to decline. The cohort used in this study only includes former dual enrollment students: first-time college students who had previously taken dual enrollment courses. These are the students who enrolled in college courses prior to fall 2014 while still in high school. Students were identified as former dual enrollment students if their enrollment or degree record prior to fall 2014 was before the student turned 18 years old. Former dual enrollment students represent 19.49 percent of the fall 2014 cohort. As a proportion of the sample for each sector, former dual enrollment students represent 23.09 percent of the students who started in four-year public institutions, 16.57 percent of the students who started in two-year public institutions and 17.20 percent of those who started in four-year private nonprofit institutions. Only 6.92 percent of the students who started in four-year private for-profit institutions had prior dual enrollments.

Enrollment Intensity

In this report, enrollment intensity is classified as exclusively full-time, exclusively part-time, or mixed enrollment (including both full-time and part-time enrollments) throughout the study period. Each of the enrollment type is based on students' enrollments across all terms in which the student was enrolled. Hence, the report does not exclude stop-outs through the first completion or, for non-completers, through the entire study period. In establishing students' enrollment intensity in this way, enrollments during summer terms (defined as terms with both the start date and the end date falling between May 1 and August 31 in any given year) and short terms (defined as terms lasting less than 21 days) were excluded from consideration.

For terms in which a student showed concurrent enrollment records (i.e., records that overlapped by 30 days or more), the two highest-intensity enrollments were considered. For example, a student concurrently enrolled half-time at two institutions was categorized as enrolled full-time for that term. In doing this, we create one single enrollment record from a set of concurrent enrollment records. The enrollment status for the single enrollment record is defined as full-time if (1) for terms with concurrent enrollments, the two highest-status enrollment records included at least one full-time enrollment, or one three quarter-time enrollment and one at least less than half-time enrollment; or (2) for terms with concurrent enrollments, the two highest-status enrollment records both reflected half-time enrollment. The enrollment status for the single enrollment record is defined as half-time, if for terms with concurrent enrollments, the two highest-status enrollment records included some combination of half-time and less than half-time enrollments, but no full-time enrollment, and no more than one half-time enrollment.

Overall, for each term under consideration (i.e., all terms except summer terms and short terms – less than 21 days – up through the first completion, or, if no completion, throughout the entire study period), the “exclusively full-time enrollment” designation was assigned to students whose enrollment showed exclusively full-time enrollment for all terms. The “part-time enrollment” designation was assigned to students whose enrollment for each term under consideration showed exclusively three quarter-time, half-time or less than half-time enrollment.

The category of mixed enrollment was applied to students who showed a combination of full-time and part-time enrollments across the terms under consideration. Finally, students who showed records indicating withdrawal (i.e., students who were enrolled, but withdrew before the term ended) but no full-time or part-time enrollments were randomly assigned to an enrollment intensity category.

Concurrent Completion

For this report, we examined completion by first-time students at either two-year or four-year institutions. We defined completion as having obtained a degree or certificate at any institution within the six-year or eight-year study period (i.e., by June 30, 2020). Clearinghouse data provide a unique headcount of U.S. college enrollments during each term, which allows for the tracking of individuals including those with concurrent completion. In preparing data for this report, a small number of individuals showed more than one completion awarded at multiple institutions on the same day. In these instances, a primary completion record was selected using decision rules specific to the sector of the student’s starting institution.

The first set of decision rules was applied to students with concurrent completions who started at a two-year institution:

1. Concurrent Completions at Two Different Two-Year Institutions
 - Same institution over different institution: Completions at the starting institution were selected over completions at other institutions.
 - Random selection: If the first decision rule did not result in a single completion record being selected, then a completion record was selected at random.
2. Concurrent Completions at a Two-Year Institution and a Four-Year Institution
 - Two-year then four-year: The two-year degree completion was considered the first completion and the four-year degree completion was considered a subsequent completion.
3. Concurrent Completions at Two or More Four-Year Institutions
 - Random selection: If a student started at a two-year institution but later completed at two or more four-year institutions concurrently, then a completion record was selected at random.

The second set of decision rules was applied to students who started at four-year institutions and later showed concurrent completion records:

1. Same institution over different institution: Completions at the starting institution were selected over completions at other institutions.
2. Four-year over two-year: If the first decision rule did not result in the selection of a single completion record, then completions at four-year institutions were selected over those at two-year institutions.
3. Random selection: If neither of the first two decision rules resulted in the selection of a single completion record, then a completion record was selected at random.

Imputation of Values for Gender

The Clearinghouse’s coverage of student gender has increased dramatically for enrollments occurring in recent years. However, imputation of gender for the majority of enrollment records is still necessary in order to use the data for research studies using older cohorts. To meet this need, 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 increase 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. The imputation process that yielded additional gender codes produced a total gender coverage rate of 94 percent.

Imputation of Missing DegreeVerify Graduation Data

The Clearinghouse collects graduation information from its participating institutions via two data reporting services: Enrollment Reporting and DegreeVerify. Enrollment Reporting has higher data coverage rates, but includes only basic completion information such as graduation indicator and the date of graduation. For the fall 2014 cohort, Enrollment Reporting covered 96 percent of all the students in Title IV degree-granting institutions listed in IPEDS (including 99 percent of the students in public institutions, 95 percent in private nonprofit institutions, and 70 percent in private for-profit institutions). DegreeVerify includes enhanced information on completions, including degree title, major, level, and CIP code, but covered only 94 percent of enrollments in 2014. Institutions may participate either in Enrollment Reporting alone or in both services. Completions data for this report included information drawn from either service.

An analysis conducted by the Clearinghouse on the 2014 cohort found that graduation data for the institutions that participated in DegreeVerify were relatively more complete for some of the years covered in this study than those for institutions that participated only in Enrollment Reporting, biasing completion rates slightly downwards for institutional sectors with lower participation rates in DegreeVerify.

To correct for this bias, the Research Center conducted a randomized imputation procedure for missing graduation data among students at non-DegreeVerify institutions who were no longer enrolled but for whom outcome data were missing (that is, for whom the institution had reported neither a graduation nor a withdrawal status in their Enrollment Reporting). This involved comparing the Enrollment Reporting and DegreeVerify records for institutions that participated in both services and estimating, for each institution type, the average percentage of students with missing outcomes in the enrollment data who had a reported graduation in the DegreeVerify data. We further specified these underreporting rates by taking into account student age and the academic year. We then used random assignment of graduation outcomes to students with missing data at the institutions that did not participate in DegreeVerify to match each institution's underreporting rate for each student age group and for every year of the study to the average rate for similar students at institutions of the same type that did participate in DegreeVerify.

This imputation was performed only for students with missing outcomes data at institutions that did not participate in DegreeVerify. It is based on the typical underreporting of graduation outcomes from similar institutions that participate in both Enrollment Reporting and DegreeVerify. The table below shows, for each institution type, the percentage of the starting cohort for whom graduation data were imputed:

Institutional Sector	Percentage of the Starting Cohort With an Imputed Completion
Private For-Profit Two-Year	0.60%
Private Nonprofit Two-Year	0.75%
Public Two-Year	0.35%
Private For-Profit Four-Year	0.41%
Private Nonprofit Four-Year	0.40%

Public Four-Year	0.16%
Total	0.28%

Adjusting Noncoverage through Weighting by State and Institution Type

The institutions participating in the Clearinghouse Enrollment Reporting service do not cover 100 percent of all Title IV, degree-granting institutions in the U.S. To account for possibilities of not capturing a student’s enrollment outcome because of non-coverage of Clearinghouse data, weights were calculated using the 2019 coverage rate of the sector, control, and state of the institution where a student was enrolled for the last enrollment record.

Our data covers three different groups of students: First, students who completed a degree at the starting institution or were still enrolled at the starting institution by the end of the study period. Second, students who completed a degree or were still enrolled at a different institution by the end of the study period. And third, students who stopped out by the end of the study period. However, the number of students who stopped out is likely overestimated due to under coverage. Specifically, students that show up as no longer enrolled may have potentially transferred to an institution that does not submit data to the Clearinghouse. Therefore, the number of stop-outs in the U.S. is likely lower than observed in our data, and the number of transfer students is likely higher. To take this into account, we overcount transfer students and undercount stop-outs (“missings”).

For students who completed a degree or were still enrolled at a different institution by the end of the study period, a “transfer” weight > 1 was applied. This transfer weight is based on the coverage of the sector, control, and state of the institution in which the student was enrolled for the last enrollment record, as calculated by the formula provided below:

$$\text{Transfer Weight} = \frac{\text{Transfer Counts} + [\text{Missing Counts} * \text{Noncoverage Rate} * \frac{\text{Transfer Counts}}{\text{Nonmissing Counts}}]}{\text{Transfer Counts}} > 1$$

For students who stopped out by the end of the study period and for whom any further observations are missing, a “missing” weight < 1 was applied. This missing weight is based on the coverage of the sector, control, and state of the institution in which the student was enrolled for the last enrollment record, as calculated by the formula provided below:

$$\text{Missing Weight} = \frac{\text{Transfer Counts} - [\text{Missing Counts} * \text{Noncoverage Rate} * \frac{\text{Transfer Counts}}{\text{Nonmissing Counts}}]}{\text{Transfer Counts}} < 1$$

Note for completeness that we assign a weight of one to the first group of students who completed a degree at the starting institution or were still enrolled at the starting institution by the end of the study period.

Data Limitations

The data limitations in this report center mainly on the data coverage, the methods used for cohort identification, and the definition of key constructs (as outlined above). The representation of private for-profit institutions in the Clearinghouse data is lower than that of other institution types, with 80 percent coverage for four-year private for-profit institutions in fall 2014 compared to 95 percent and 99 percent respectively for four-year private nonprofit institutions and four-year public institutions. Despite the challenges presented by low participation in the early years covered in this report, current Clearinghouse data nevertheless offer near-census national coverage, representing 97 percent of U.S. postsecondary

enrollments. In an effort to correct for coverage gaps in this study, data were weighted (as explained above).

Data limitations resulting from the cohort identification methods used in preparing this report should also be noted. Because the Clearinghouse data on designations for class year are incomplete, the researchers identified first-time undergraduate students via two indirect measures:

- No previous college enrollments recorded in the Clearinghouse data, and
- No previous degree awarded in the Clearinghouse's historical DegreeVerify database.

Given these selection criteria, the sample for this report may include students who had more than 30 Advanced Placement (AP) or International Baccalaureate (IB) credits. It is also possible that a small number of graduate students are included in the study cohort because of inconsistencies in the historical depth of DegreeVerify database records.

Finally, although Clearinghouse data contain some demographic information on students, historical coverage rates for the demographic data elements are uneven. Consequently, results on gender are based partially on imputed values, as described above and one in five students in the cohort had the race/ethnicity value unknown or missing. No imputation was attempted for missing race/ethnicity data.

Suggested Citation

Causey, J., Huie, F., Lang, R. Ryu, M., & Shapiro, D.. (December 2020), *Completing College 2020: A National View of Student Completion Rates for 2014 Entering Cohort (Signature Report 20)*, Herndon, VA: National Student Clearinghouse Research Center.