

**College Scorecard Data by Field of Study**

DRAFT as of 10/4/19

Contents

[Introduction 2](#_Toc16782809)

[Data Methodology 2](#_Toc16782810)

[Field of Study identifiers 2](#_Toc16782811)

[Post-completion earnings 3](#_Toc16782812)

[Cumulative federal loan debt 4](#_Toc16782813)

[Program size 5](#_Toc16782814)

[Data Sources and Data Limitations 5](#_Toc16782815)

[Federally aided students 6](#_Toc16782816)

[Aggregation of Data 6](#_Toc16782817)

[Earnings 7](#_Toc16782818)

[CIP code granularity 7](#_Toc16782819)

[Data quality 7](#_Toc16782820)

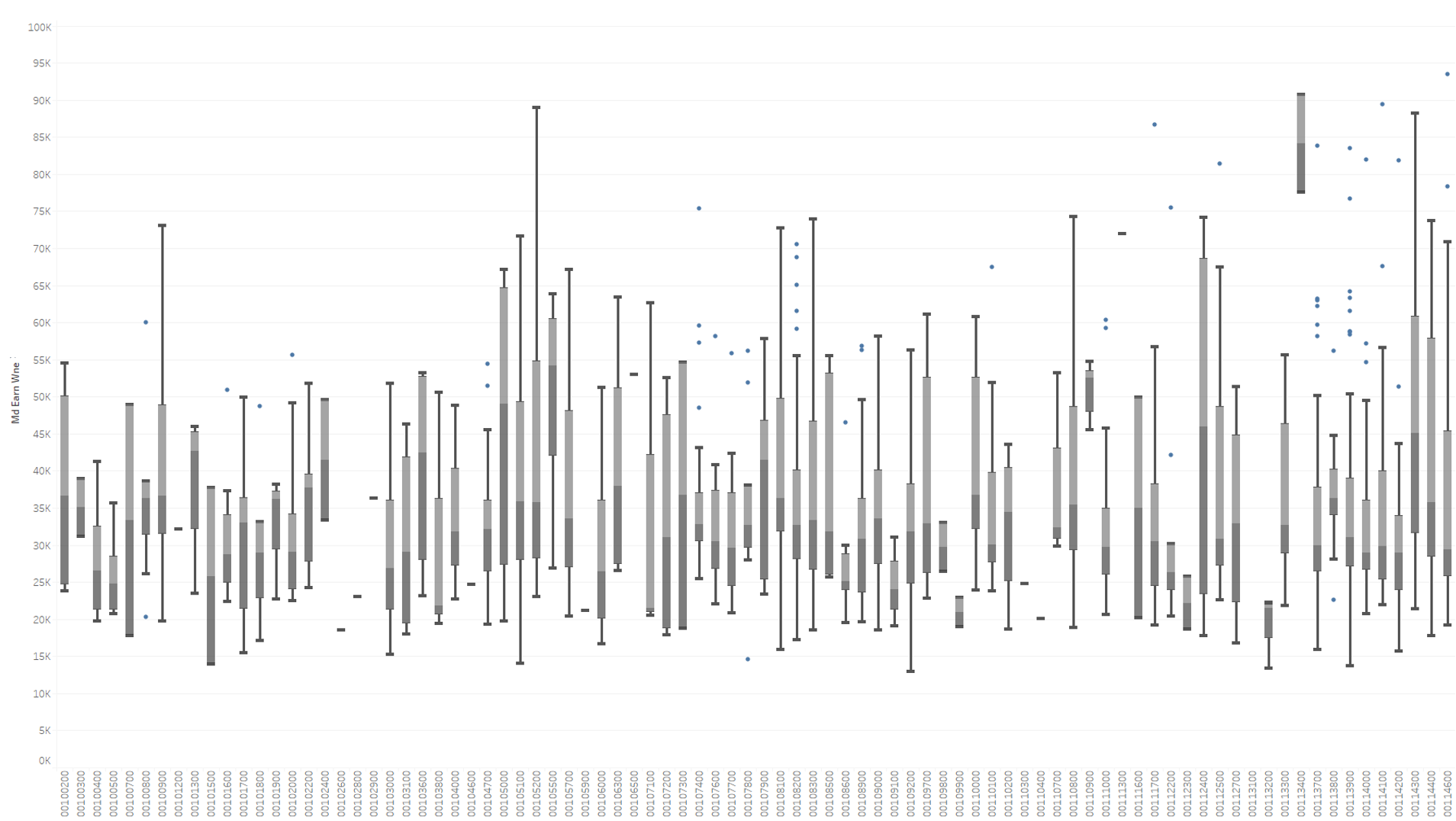
[Considerations and Potential Risks: 8](#_Toc16782821)

# Introduction

As part of the implementing the [Improving Free Inquiry, Transparency, and Accountability at Colleges and Universities](https://www.whitehouse.gov/presidential-actions/executive-order-improving-free-inquiry-transparency-accountability-colleges-universities/) Executive Order, the College Scorecard (Scorecard) expanded to include new data elements describing post-graduate earnings and cumulative loan debt of graduates by field of study. The primary purpose of the Scorecard is to provide data to help prospective postsecondary students make informed enrollment decisions.

Since 2015, Scorecard has provided information describing institutions as a whole. However, depending on the field of study, outcomes can vary substantially within institutions.

Figure X. Median earnings for undergraduate fields of study, by institution .



Note: only the first 85 institution boxplots are shown as examples of earnings variation across fields of study within institutions.

In order to provide more comprehensive, granular data to prospective students, Scorecard is providing information on post-graduate earnings and cumulative debt by field of study. The purpose of this document is to describe the data generating process and data limitations of data elements disaggregated by field of study. For detailed data documentation describing the existing institution-level data metrics, see the [Scorecard documentation report](https://collegescorecard.ed.gov/assets/FullDataDocumentation.pdf).

# Data Methodology

The methodology used to calculate earnings and cumulative debt by field of study were based, in part, on discussions from the spring 2019 Scorecard [Technical Review Panel](https://edsurveys.rti.org/cs_trp/index.aspx). The [Scorecard consumer-facing website](https://collegescorecard.ed.gov/assets/FullDataDocumentation.pdf) includes only a subset of data elements (e.g. median earnings, median cumulative debt) disaggregated by field of study for currently operating undergraduate institutions[[1]](#footnote-1) and a more comprehensive set of data elements, including counts and averages are provided in the Scorecard downloadable data files and application programming interface (API).[[2]](#footnote-2) Calculations are made for fields of study in all institutions that participate in the Federal student financial assistance programs administered by the Department under Title IV of the Higher Education Act of 1965 as amended. This document describes the data methodology for all data calculated by field of study.

## 

## Defining Field of Study

The unit of analysis for calculations by field of study is defined as the unique combination of an institution (six-digit OPEID), a four-digit Classification of Instruction Program (CIP) code, and a credential level.

* Institutions are defined as having a unique office of postsecondary education ID number (OPEID). For purposes of Scorecard calculations, this is based on the six-digit OPEID (**OPEID6)**, representing the main campus (and all of its branches) of an institution. Scorecard uses this aggregated approach because not all institutions report information needed to calculate these metrics at the more granular branch location level. While data are calculated at the OPEID6-level, they are disseminated at the Integrated Postsecondary Education Data System (IPEDS) UNITID-level to allow fields of study not available at particular branches of institutions to be excluded from consumer-facing tools. The OPEID6-level medians and averages are repeated across any branches that offer those fields of study based on IPEDS Completions component reporting. All fields of study are displayed under the main campus UNITID.
* CIP codes are a taxonomy of academic disciplines classified by a six-digit code with the first two digits representing the most broad category for a discipline, the first four digits representing a more granular category, and all six digits represent the most granular description of an academic discipline. Scorecard uses the first four digits of the CIP code in its calculations (**CIPCODE)**.because four-digit CIP code descriptions are easier for prospective students to understand and because combining six-digit CIP codes leads to larger cell sizes, which in turn lead to fewer data points that need to be privacy suppressed.
* The credential level (**CREDLEV)** codes correspond to the following categories of programs:
  + 1: Undergraduate Certificates or Diplomas
  + 2: Associate's Degrees
  + 3: Bachelor's Degrees
  + 4: Post-baccalaureate Certificates
  + 5: Master's Degrees
  + 6: Doctoral Degrees
  + 7: First Professional Degrees
  + 8: Graduate / Professional Certificates
  + 99: Non-credential program

## Program size

In order to give prospective students a sense of program size, Scorecard provides the number of awards conferred by field of study (IPEDSCOUNT1 and IPEDSCOUNT2) as reported by institutions to IPEDS. This count of awards represents all awards, including first major awards, second major awards, etc. That is, a student who completed multiple awards would be measured multiple times in these counts. IPEDS award levels do not exactly match the credential levels in the National Student Loan Data System (NSLDS), the source of cohorts included in the field of study data. In order to provide IPEDS counts that correspond to the NLSDS credential levels, some aggregation of IPEDS award levels was needed. Figure X below indicates the recoding and aggregation scheme applied to IPEDS to calculate the number of awards provided in the field of study data.

Figure X. NSLDS credential level to IPEDS award level correspondence.

|  |  |
| --- | --- |
| **NSLDS Credential Level** | **IPEDS Completions Data Award Levels Aggregated** |
| 1: Undergraduate Certificates or Diplomas | Less-than-1-year At least 1- but less-than-2-year At least 2- but less-than-4-year certificates |
| 2: Associate's Degrees | Associate's degrees |
| 3: Bachelor's Degrees | Bachelor's degrees |
| 4: Post-baccalaureate Certificates | Not available |
| 5: Master's Degrees | Master's degrees |
| 6: Doctoral Degrees | Doctor's degrees-research/scholarship Doctor's degrees-other |
| 7: First Professional Degrees | Doctor's degrees - professional practice |
| 8: Graduate / Professional Certificates | Post-baccalaureate certificatesPost-master's certificates |
| 99: Non-credential program | Not available |

## Post-completion earnings

While the long-term goal is to provide earnings data points for students several years after completion, the Department only recently began collecting information by field of study. In the short term, Scorecard will calculate based on what’s available. As of 2019, Scorecard provides one-year post-completion earnings by field of study (MD\_EARN\_WNE) for the cohort of federally aided students who completed in award years 2014-15 and 2015-16 with earnings measured in calendar years 2016 and 2017 respectively (given in constant 2018 dollars). The cohort of evaluated students excludes those who

* were subsequently enrolled in school during the measurement year,
* had died prior to the beginning of the measurement year,
* who had loan deferments for military service during the measurement year,
* and did not work during the measurement year.

In addition, Scorecard provides the

* count of graduates measured for each earnings value (TITLEIVCOUNT), excluding students due to enrollment, military service, or death
* the count of graduates who were working in the measurement year, whose earnings are included in the calculated median value (EARNINGSCOUNT).

Students who completed multiple awards (e.g. double major) were measured multiple times only if multiple awards were completed in a different four-digit CIP code, and/or credential level. Students who completed multiple awards within the same institution at the same four-digit CIP code and credential level were measured only once.

The earnings measurement represents the sum of wages and deferred compensation from all non-duplicate W-2 forms and positive self-employment earnings from Form 1040 Schedules SE (Self-Employment Tax) for each student measured.

## Cumulative federal loan debt

Scorecard provides the mean (DEBTMEAN) and median (DEBTMEDIAN) federal loan debt of federal borrowers by field of study for cohorts of students who complete in the same two-year timeframe. In addition, Scorecard provides the borrower count associated with each mean and median federal loan debt value (COUNT).

The cumulative federal loan debt includes only loan disbursement amounts and does not capture any accrued interest, even if that interest accrued prior to graduation. The cumulative loan debt only includes loans disbursed at the same academic level (i.e. graduate, undergraduate) as the evaluated credential level. For example, calculated debt values for a Master's degree-level field of study will include all graduate-level loans and exclude all undergraduate-level loans. Academic levels are assigned to credential levels based on the following categorizations:

* Undergraduate credentials:
  + Undergraduate Certificates or Diplomas
  + Associate's Degrees
  + Bachelor's Degrees
  + Post-baccalaureate Certificates
* Graduate credentials
  + Master's Degrees
  + Doctoral Degrees
  + First Professional Degrees
  + Graduate / Professional Certificates

The cumulative federal loan debt includes all debt (at the same academic level of the evaluated credential level) dispersed by the evaluated institution regardless of whether the evaluated student switched fields of study at the evaluated institution. For example, if a student began as an undergraduate general studies student and then earned a bachelor’s degree in agriculture at the same institution, the debt calculated for the agriculture bachelor’s degree would include debt used to enroll in the prior general studies field of study.

All Stafford and Graduate PLUS loans are included in the sum of cumulative federal debt. Cumulative federal debt values exclude parent PLUS loans and Perkins loans.

Students who completed multiple awards (e.g. double major) were measured multiple times only if multiple awards were completed in a different four-digit CIP code, and/or credential level. Students who completed multiple awards within the same institution at the same four-digit CIP code and credential level were measured only once.

## 

## Privacy protection

Data presented by field of study have been treated with cell suppression methods to reduce the risk of disclosure. To further reduce disclosure risk, information about the specific suppression rules is not available to the public. Data points suppressed for privacy are indicated by the “PrivacySuppressed” data code. While the majority of the fields of study offered by institutions have privacy suppressed data values for cumulative loan debt and post-completion earnings, the fields of study for which there are unsuppressed (reportable) data values represent a large share of students who graduate.

Figure X. Proportion of data that is reportable following privacy suppression

|  |  |  |  |
| --- | --- | --- | --- |
| **Metric subject to privacy suppression** | **UNITIDs with at least 1 reportable field of study** | **Reportable UNITID-by-field-of-study combinations** | **IPEDS awards from UNITIDs with reportable values** |
| Median debt | 76.1% | 24.3% | 68.7% |
| Median earnings | 78.6% | 21.4% | 66.3% |
| Median debt and median earnings | 72.0% | 19.4% | 62.8% |

# Data Sources and Limitations

The National Student Loan Data System (NSLDS) is the Department of Education’s (Department) central database for monitoring title IV federal student aid with data on federal borrowers and grant recipients dating back to the 1960s. While primarily used for operational purposes, such as tracking federal grant disbursements and loan disbursements and repayment, NSLDS provides administrative data from which loan debt data elements and earnings cohort construction can be derived.

To gain insight into the labor market outcomes of individuals who graduate, Scorecard linked NSLDS records to administrative tax records maintained by the IRS within the Department of the Treasury. Specifically, post-completion earnings values are derived from the sum of wages and deferred compensation from all W-2 forms received for each individual, plus self-employment earnings from Schedule SE.

The Integrated Postsecondary Education Data System (IPEDS) is a system of surveys conducted annually by the National Center for Education Statistics (NCES). The completion of all IPEDS surveys is mandatory for all institutions that participate in, or are applicants for participation in, any federal financial assistance program authorized by Title IV of the Higher Education Act of 1965, as amended. Scorecard derives the count of awards conferred by field of study from the IPEDS completion survey results. In this survey, institutions report all awards conferred by award level and six-digit CIP code.

## Federally aided students

Since Scorecard earnings and cumulative loan debt data by field of study are derived from NSLDS data, these data elements describe only those students who received federal financial aid in the form of title IV grants and loans. Given not all graduates received title IV aid to attend school, data users should not assume that Scorecard data values are representative of all students who graduated from a particular field of study.

Scorecard provides cumulative debt calculations based on only federal loans for only title IV borrowers. Further, Scorecard does not capture the amount borrowed from non-federal sources such as loans administered by states, institutions of higher education, or other private entities. As an example of the prevalence of nonfederal borrowing, 15 percent of undergraduate fourth-year seniors, between the ages of 18 and 24, had nonfederal loans in 2015-16.[[3]](#footnote-3) In addition, debt calculations are based only on the amount borrowed and does not include capitalized interest, even if that interest accrued prior to graduation. Given these factors, Scorecard suggests users interpret these data with the understanding estimates may often be lower than the amount owed upon graduation.

Earnings are estimated for only those who receive title IV federal financial aid. Figure X examines the prevalence of title IV recipients among undergraduates in 2015-16. This figure shows that, nationally, 54 percent of undergraduate students received any form of title IV aid and the prevalence of title IV recipients varies across different types of institutions.

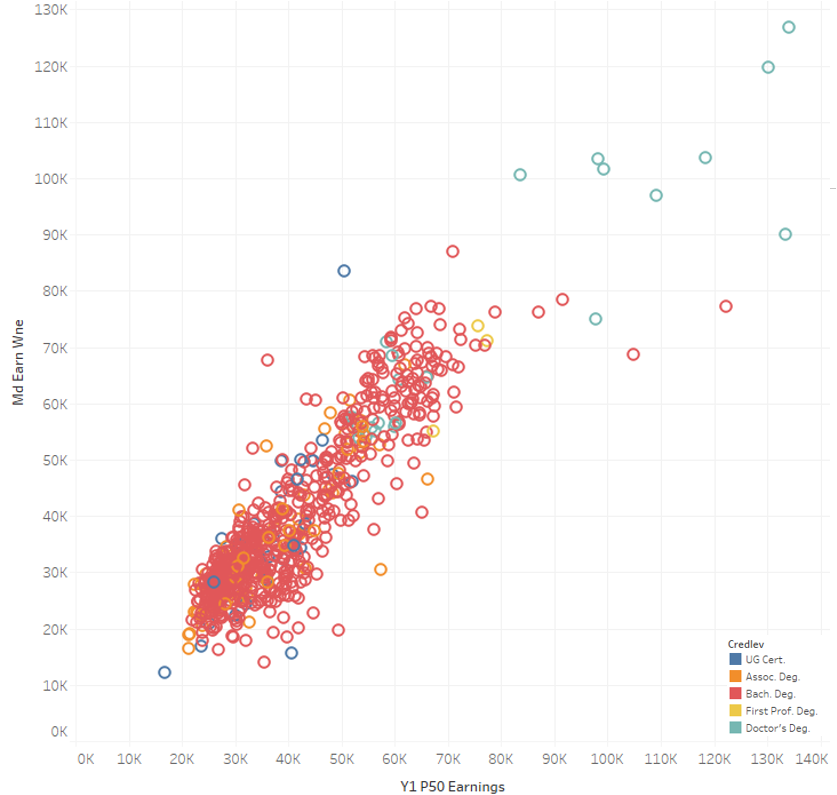
|  |  |  |
| --- | --- | --- |
| **Figure X. Percentage of undergraduates receiving federal Title IV aid, by control and level of institution, 2015–16** | | |
| **All undergraduates** | | **54.4** |
| Control | Level |  |
| Public | Less-than-2-year | 42.8 |
| 2-year | 38.8 |
| 4-year | 59 |
| Non-doctorate-granting | 53 |
| Primarily subbaccalaureate3 | 42.5 |
| Primarily baccalaureate | 62.2 |
| Doctorate-granting | 61.8 |
| Private | Less-than-4-year | 74.5 |
| 4-year | 63.4 |
| Non-doctorate-granting | 67.1 |
| Doctorate-granting | 60.7 |
| Proprietary | Less-than-2-year | 77.6 |
| 2-year | 75.6 |
| 4-year | 76.7 |

Source: <https://nces.ed.gov/pubs2018/2018466.pdf>

The limitation of title IV recipients also may affect cohort construction. Students are excluded from the earnings measurement if they are subsequently enrolled in school during the earnings measurement year. For example, if a student who received a BA in 2015-16 and was subsequently enrolled in calendar year 2017 for an MA program, that student would be excluded from the earnings calculation. Scorecard will not be able to correctly capture the subsequent enrollment status for this type of student if he or she does not receive either title IV aid in the MA program or an in-school deferment for loans previously taken out.

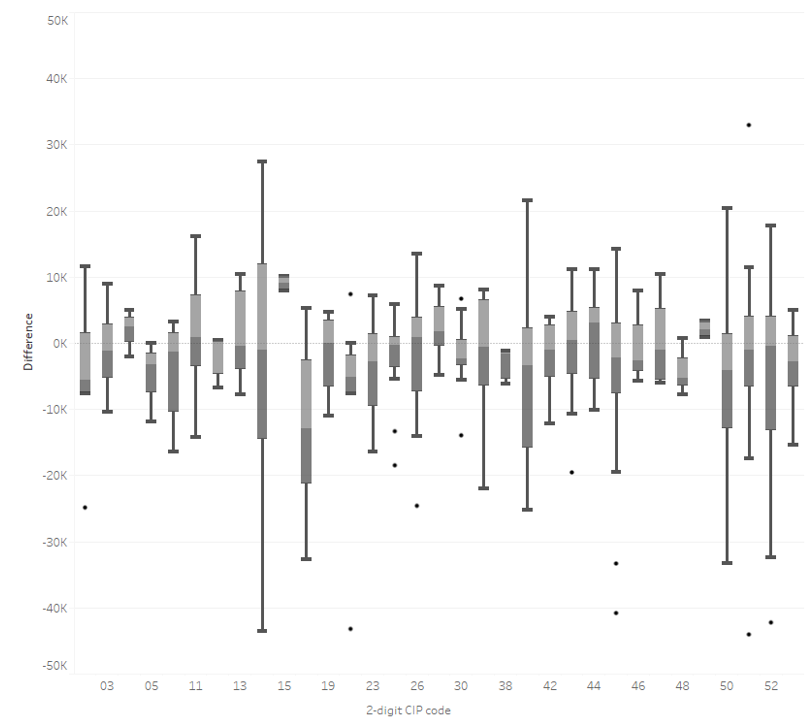
Scorecard earnings data were compared to publicly available data from the Postsecondary Employment Outcomes (PSEO) data from the U.S. Census Bureau. Institutions and fields of study were matched between the sources and the 1 year post-completion earnings were analyzed to confirm general agreement between the data sources. Due to methodological differences (e.g., the PSEO does not restrict the cohort to federally-aided students; the PSEO uses unemployment insurance wage records rather than IRS tax records), exact matching was not expected. It is not clear if these differences would be expected to lead to higher or lower median earnings when comparing Scorecard to PSEO data. However, when examining the scatter plot provided in Figure X, the positive correlation between the Scorecard and PSEO data suggests that the two sources are generally in agreement.

Figure X. Scatterplot of Scorecard median earnings and median earnings from Census PSEO earnings data.



Despite the positive correlation, there is still considerable variation by field of study. In figure X below, the fields of study are grouped into their broad areas of study (2-digit CIP codes), though the differences shown are still comparisons at the field of study level. Some broad areas of study seem more variable than others, and it may be, for example, that the fields of study contained in those areas show the greatest cohort characteristic differences between Scorecard and PSEO cohorts. Without the underlying data we cannot be certain. Other areas of study show very little difference between the PSEO and Scorecard earnings values for the associated fields of study. In all, consider all field of study differences regardless of the CIP code grouping, the median difference was $-1,983, with an interquartile range from $-11,805 to $4,695, which is not unreasonable given the methodological differences.

Figure X. Differences in median earnings comparing Scorecard to U.S. Census PSEO data



## Aggregation of Data

CIP codes are a taxonomy of academic disciplines developed where academic disciplines are classified by a six-digit code with the first two digits representing the most broad category for a discipline, the first four digits representing a more granular category, and all six digits represent the most granular description of an academic discipline. Even within the most granular category of discipline, institutions often provide different types of program offerings with a different course catalog program name. Scorecard uses the first four digits of the CIP code in its calculations. In doing so, Scorecard can provide information in categories that may be easier to understand for prospective students. In addition, Scorecard can provide more information that does not privacy suppressed because combining numbers across sub- categories increases cell sizes. The trade-off is lack of granularity in describing individual program offerings by institutions. For example, the following six digit CIP codes would all be aggregated into one calculation for CIP code 5004, *Design and Applied Arts.*

* 500401: Design and Visual Communications, General
* 500402: Commercial and Advertising Art
* 500404: Industrial Design
* 500406: Commercial Photography
* 500407: Fashion/Apparel Design
* 500408: Interior Design
* 500409: Graphic Design
* 500410: Illustration
* 500411: Game and Interactive Media Design
* 500499: Design and Applied Arts, Other

To the extent that outcomes vary across these sub- categories, this variation would not be observed using the current methodology.

The Department recognizes separate branch “locations” for all Title IV-eligible institutions that have a Program Participation Agreement (PPA), the formal document establishing their eligibility to participate in Title IV programs, and assigns each location an 8-digit OPE identification number (OPEID). Similar to CIP code construction, the OPEID that identifies individual institutions and their branch locations is hierarchal. Related institutions use the same first six digits of an OPEID, which is tied to the “main campus”. The last two digits identify specific branch locations in a larger family of campuses. While NSLDS is increasingly moving towards reporting student enrollments and aid-receipt at the 8-digit OPEID level, this is not yet universal and many institutions report information on where students enroll or receive aid only at the 6-digit OPEID level (since the main campus is often where aid is managed for all branches of an institution). Given the lack of consistency with reporting to NSLDS at the more granular 8-digit-OPEID-level, Scorecard constructs cohorts and calculates data values by field of study using the 6-digit OPEID institutional identifier. To the extent that outcomes vary across branch locations, this variation would not be observed using the current methodology.

## First-year earnings

Scorecard’s long-term goal is to measure the annual earnings of former students at multiple points in time after graduation up to 10-years post-graduation. However, NSLDS did not begin collecting data at the program-level until the 2014-15 award year. The lack of program-level data prior to 2014-15 limits historical cohorts that can be derived from NSLDS by field of study. Further, Scorecard combines students into two-year cohorts to maximize cohort cell sizes in order to minimize privacy suppressions. Therefore, first complete two-year cohort that can be derived from NSLDS by field of study is the 2014-15 and 2015-16 award years. With 2017 calendar year (and prior) earnings available, Scorecard provides first-year earnings for those who graduate in 2014-15 and 2015-16 measured in calendar years 2016 and 2017 respectively. As more data become available, Scorecard will calculate salary after graduation measured multiple years post-graduation. Data users should use caution in using these first-year earnings data to predict longer-term earnings outcomes because research suggests a week association between earnings measured soon after graduation versus earnings measured many years later. For example, a recent publication by the National Bureau of Economic Research indicated that first year and seventh year metrics are barely correlated, suggesting labor outcome data may be misleading if measured too soon after graduation.[[4]](#footnote-4)

## Nonemployment

Earnings data include information on the fraction of individuals who did not work for pay among those who are not currently enrolled. This is based on information about the number of individuals with no reported earnings over the course of the full year. Data users should be careful in interpreting this as a measure of unemployment, meaning the fraction of workers in the labor force (actively searching for a job) who are unable to find employment. In particular, data users should note that those who did not work may have chosen to be out of the labor market for a variety of reasons.

## Data quality

In 2014-15, FSA began requiring institutions to report program-level enrollment data to NSLDS. While data reporting seems to have stabilized over time, enrollment reporting in the earlier years was likely incomplete and potentially inaccurate. In order to address this, Scorecard published preliminary cumulative loan debt data in May 2019 and the Department asked institutions to make corrections to their own historical NSLDS program-level enrollment data by mid-July 2019. In order to examine the overall data quality of program-level reporting to NSLDS after institutions were able to make historical updates and corrections, the Department compared NSLDS program-level data with other internal Federal Student Aid (FSA) reporting and available samples from NCES statistical collections.

To assess consistency of reporting by institutions, the Department examined the process where institutions initially report program-level CIP codes at the time of aid disbursement and then reconfirm or update those CIP codes for students via NSLDS program-level enrollment reporting; for the years in question, those CIP codes are almost always either confirmed or corrected during that secondary FSA process step, which was observed at 97% for the 2014-15 award year data.

To assess the completeness of reporting by institutions, institution-level graduation events in NSLDS can serve as a comparison. Institution-level graduation events serve as a trigger for a student entering repayment due to a graduation and should have a corresponding program-level graduation rate. Figure X provides aggregate ratios of program-level graduation events to institution-level graduation events if they occur within 180 days for an individual student at an individual institution (that is, for every institutional graduation event, how often is there a corresponding program-level graduation event). The table shows that the 2014-15 reporting was less complete than in subsequent years.

|  |  |
| --- | --- |
| **Figure X. Percentage of IHE -level graduation events with corresponding program-level graduation events as of mid-July 2019** | |
| **Award year** | **Percentage** |
| 2014-15 | 73% |
| 2015-16 | 90% |
| 2016-17 | 91% |
| 2017-18 | 92% |

Source: NSLDS as of July 2019

Further analysis suggests that these ratios are not substantially different, in the aggregate, across institutional characteristics including institutional control (i.e. Public, Private, Proprietary) and institution size (i.e. number of completers) and also across groups of students who received grants and loans versus those who received only grants.[[5]](#footnote-5)

Further examination of NSLDS data (Figure 2) shows that in the initial year of reporting (2014-15) the typical institution reported 80 percent of institutional graduation events were also reported at the program-level, with 25 percent of institutions reporting less than 45 percent of institutional graduation events were also reported at the program-level. In the next three years of reporting, the typical institution reported at 93, 95, and 95 percent of institutional graduation events were also reported at the program-level, respectively.

**Figure 2. Distribution of Ratios for Institution-Level Graduation Events vs Program-Level Graduation Events**



Source: NSLDS as of July 2019

<< insert NCES sample comparison information /analysis here >>

Debt and earnings metrics can also be biased if institutions reported incorrect CIP codes to NSLDS (for example, incorrectly reporting nursing majors as business majors). NCES researched program-level graduation event data in NSLDS for respondents from the 2012/17 Beginning Postsecondary Students Longitudinal Study (BPS:12/17), comparing their enrollment information in NSLDS to data from the National Student Clearinghouse (NSC) and the student survey component of BPS:12/17. BPS contains a sample of first-time beginning students enrolled in multiple types of institutions and degree programs for six years (2012-2017). Figure 3 further limits the BPS sample members to federal aided students.

Figure 3 shows the percentage of cases that match out of all aided sample members that had a graduation event reported by either of the two comparison data sources. Matches represent an indication that graduation events were observed in both data sources; *‘un-matched’ events from either source could be due to a lack of records in the second data source OR a difference in records both present and observed in each of the two data sources*.

Figure 3 match rates are disaggregated by credential level. Given the sample is based on a 2011-12 beginning postsecondary sample and graduation events are evaluated beginning in 2014-15, undergraduate certificates and associate degree students have small sample sizes and their match rates should be interpreted with caution.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Figure 3. BPS:12/17 Program-Level Completions Between 2015-2017: Match Rates Within Federally Aided1 Students, Matched Student School Pairs, and Including 2-Digit General CIP Codes, by Academic Year and Credential Level** | | | | | | | | |
|  | NSC & NSLDS  pooled completions/2-digit CIP codes | |  | BPS interview & NSLDS pooled completions/2-digit CIP codes | |  | NSC & BPS interview pooled completions/2-digit CIP codes | |
| Academic Year | Number | Percent Matched |  | Number | Percent  Matched |  | Number | Percent  Matched |
| **Overall** |  |  |  |  |  |  |  |  |
| 2014-15 | 3,270 | 59.9 |  | 3,370 | 18.5 |  | 3,320 | 19.2 |
| 2015-16 | 2,410 | 55.1 |  | 2,470 | 26.5 |  | 2,330 | 26.6 |
| 2016-17 | 1,160 | 56.6 |  | 1,210 | 22.1 |  | 1,180 | 21.3 |
| **UG2 Certificate** |  |  |  |  |  |  |  |  |
| 2014-15 | 180 | 13.6 |  | 220 | 10.0 |  | 230 | 2.7 |
| 2015-16 | 150 | 19.9 |  | 220 | 12.6 |  | 190 | 6.4 |
| 2016-17 | 100 | 22.8 |  | 130 | 20.3 |  | 130 | 4.8 |
| **Associate’s Degree** |  |  |  |  |  |  |  |  |
| 2014-15 | 410 | 39.4 |  | 440 | 10.8 |  | 530 | 10.1 |
| 2015-16 | 420 | 42.4 |  | 420 | 16.1 |  | 420 | 18.5 |
| 2016-17 | 270 | 48.4 |  | 280 | 12.0 |  | 280 | 12.8 |
| **Bachelor’s Degree** |  |  |  |  |  |  |  |  |
| 2014-15 | 2,680 | 66.1 |  | 2,720 | 20.4 |  | 2,570 | 22.5 |
| 2015-16 | 1,840 | 60.8 |  | 1,830 | 30.6 |  | 1,720 | 30.8 |
| 2016-17 | 790 | 63.7 |  | 800 | 25.9 |  | 770 | 27.1 |
|  |  |  |  |  |  |  |  |  |

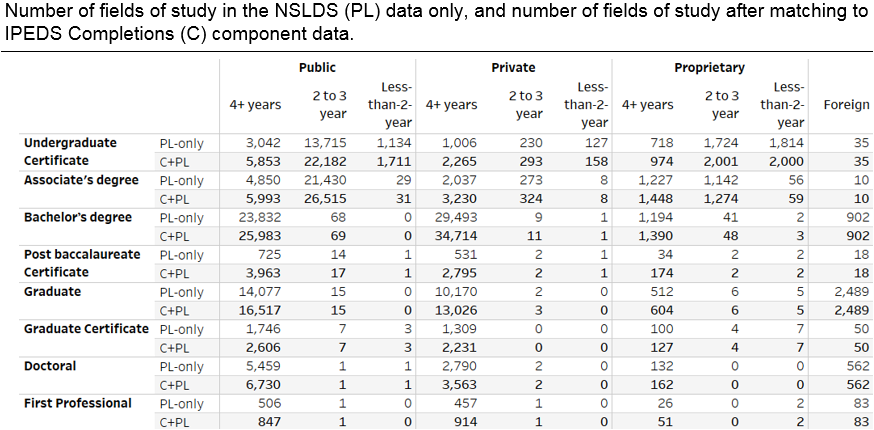
## Matching fields of study reported to different data systems

NSLDS-identified graduates should also appear in the IPEDS Completions component for the corresponding cohort (award year) period. However, the NSLDS and IPEDS classifications of students may vary due to definitional differences between the two systems. To align the two data sources, IPEDS Completions data values were aggregated across award levels to correspond to the NSLDS credential levels (as noted in Figure X). This mapping, however, is subject to bias including, but not limited to, the following situations:

* The IPEDS Post-baccalaureate Certificate category may contain some awards that NSLDS classifies as Post-baccalaureate Certificates and others that NSLDS classifies as Graduate/Professional Certificates. Examining the data, the reported fields of study in IPEDS more closely aligned with the Graduate/Professional Certificate classification.
* The NSLDS Post-baccalaureate Certificates category may contain awards that IPEDS classifies as undergraduate certificates (e.g., teacher licensure programs). Examining the data, there was no support for mapping IPEDS undergraduate certificates to the NSLDS Post-baccalaureate Certificate category.
* IPEDS has not collected a First-professional Degree category since the 2009-10 collection. These degrees are primarily reported in the Doctor’s Degree – Professional Practice or Doctor’s Degree – Other categories, but ultimately classification of formerly first-professional degrees up to the institution, and the two doctoral degree categories mentioned may contain awards that were not previously regarded as first-professional. Additionally, some non-doctoral-level programs are considered first-professional (e.g., Master’s of Divinity). As a result, there is some bias in the matching of these two IPEDS categories to the NSLDS First-professional Degree category.

While CIP code definitions do not vary across systems, there is some evidence that institutions may report graduates under different CIP codes. Figure X shows the increase in the number of fields of study before and after adding fields of study that exist only in IPEDS. These counts include non-Title IV fields of study and those where no borrowers completed, which would explain some of the increase. However, the marked increase, especially at the Undergraduate Certificate level, suggests there is additional mismatching between the two data sources.

Figure X.



Source: NSLDS and IPEDS as of March 2019

1. See the [Scorecard documentation report](https://collegescorecard.ed.gov/assets/FullDataDocumentation.pdf) for more information on the universe of institutions provided on the consumer facing tool. [↑](#footnote-ref-1)
2. See https://collegescorecard.ed.gov/data/ [↑](#footnote-ref-2)
3. See <https://nces.ed.gov/Datalab/TablesLibrary/TableDetails/12639?keyword=nonfederal&rst=true> [↑](#footnote-ref-3)
4. See <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2880319> [↑](#footnote-ref-4)
5. Analysis not shown in this document. [↑](#footnote-ref-5)