2022 Program Performance Data Description

This document describes both the 2022 Program Performance Data ("2022 PPD") file, containing the program-level data analyzed in the Notice of Proposed Rulemaking (NPRM) Regulatory Impact Analysis (RIA) for the proposed rule titled "Financial Value Transparency and Gainful Employment (GE), Financial Responsibility, Administrative Capability, Certification Procedures, Ability to Benefit (ATB)" (proposed rule), as well as a public version of these data ("Public 2022 PPD"), which excludes or suppresses some variables due to necessary privacy protections.

Details about the methodology for constructing the underlying 2022 PPD data are provided in the "Data Used in this RIA" section of the proposed rule. This document describes the unit of analysis, sample, limitations, and definitions of variables. Further detail on the sources and coding for individual variables can be found in the accompanying codebook, NPRM_2022PPD_Codebook. Due to limitations in data available to the Department of Education (ED or the Department), the methodology used to produce these data differs from the Department's proposed methodology for the Financial Value Transparency and Gainful Employment regulations in the NPRM, which relies on data elements that institutions would need to, but do not currently, report. ED believes the 2022 PPD give the best possible depiction of the rule's impact given the data currently available to the Department.

Details on privacy protections put in place in order to release the Public 2022 PPD, and how the Public 2022 PPD differs from the 2022 PPD, are also described below. As described in the "Data Used in this RIA" section of the proposed rule, the Department has re-examined the analysis presented in the RIA using the Public 2022 PPD. The Department would not have reached different conclusions on the impact of the proposed rule if it had instead relied on this privacy-protected public dataset, though the Department views analysis based on the 2022 PPD and described in the RIA to provide a more precise representation of such impact. We view the differences in the analyses as substantively minor for purposes of this rulemaking.

Overview of 2022 PPD

The unit of analysis for the 2022 PPD file is the unique combination of institutional id (opeid6), credential level (cred_lev), and classification of instructional program code at a four-digit level (cip4). The universe of programs includes all Title IV, HEA aid programs with at least one completer reported to the National Student Loan Data System (NSLDS) in award years (AY) 2015, 2016, or 2017, and that were operating as of March 2022. The program universe includes those in foreign institutions and those operating in U.S. territories.

The 2022 PPD includes information for 155,582 programs that account for more than 19 million Title IV, HEA enrollments annually in award years 2016 and 2017. This includes about 2,931,000 annual enrollments in 32,058 Gainful Employment (GE) Programs (certificate programs at all institution types, and degree programs at proprietary institutions) and about 16,337,000 annual enrollments in 123,524

non-GE Programs (degree programs at public and private not-for-profit institutions).[1] Since the sample of programs that both had completers in award years 2015-2017 and was in operation as of March 2022 is a subset of all programs, the Department estimates that programs in the 2022 PPD account for 83 percent of all Pell Grant dollars and 92 percent of Federal student loan dollars disbursed in 2016.

Consistent with the proposed regulations, the Department computed and reported debt-to-earnings (D/E) and earnings premium (EP) metrics in the 2022 PPD only for those programs with 30 or more students who completed the program during the applicable two-year cohort period for each metric—that is, those programs that met the minimum cohort size requirements.

The 2022 PPD includes five types of information assembled from various sources.

- **Program and institution information**. Identifying information about each program and institution, such as whether it is a GE program, institutional control, Classification of Instructional Programs (CIP) code, credential level, predominant award type of the institution, location, special institution status, and instructional spending per student;
- **Performance metrics**. D/E and EP metrics and the information used to compute them, including median Federal student loan debt and earnings of program completers and earnings thresholds based on the median earnings of high school graduates in the labor force and in the same state, as measured in the American Community Survey. The 2022 PPD also includes flags indicating whether the program has passing or failing D/E and EP metrics;
- **Enrollment, completions, and aid volume**. Title IV enrollment and number of completers; Pell Grant and Federal student loan volume;
- **Borrower outcomes**. Borrower outcomes of program participants, including information about default, debt service, and projected lifetime loan repayment ratios;
- **Student characteristics**. Share of all program completers (not just Title IV recipients) by race/ethnicity from the Integrated Postsecondary Education Data System (IPEDS). Though not included in the public release, the 2022 PPD includes pre-program student characteristics for Title IV completers from FAFSA, including first-generation college student status, out-of-state, family income, expected family contribution (EFC), sex, age, dependency status, and pre-program income.

Due to data limitations, the Department calculated the median earnings data for a different cohort of program completers than the one used to calculate median debt. Earnings were measured for Title IV students completing their credential in either award years 2014-2015 or 2015-2016. Students in most programs would thus have their earnings captured in the third calendar year following the end of the award year in which they completed—calendar years 2018 and 2019, respectively. Median cumulative debt was measured for completers, including non-borrowers, in the 2015-2016 or 2016-17 award year. The Department measured their debt levels in 2016 dollars (adjusting by the CPI) and used this as an estimate for the debt levels of the AY2015 and AY2016 completers assuming debt levels remained unchanged in real terms between these cohorts. As described more completely in the RIA, this approach

---

[1] The enrollment totals reported here differ slightly from those in the Public 2022 PPD due to the privacy protocols described below. Furthermore, students that attend multiple programs during the same award year, whether at the same institution or at multiple institutions, will be counted multiple times in these enrollment counts.

provides the most recent and complete program-level debt and earnings data available to the Department.

Differences from Prior Released Data and Proposed Methodology

The Department released the GE Information Rates Data at Negotiated Rulemaking on April 20, 2022. The 2022 PPD differs in several ways from this data. Briefly, these differences are:

- 2022 PPD provides median debt, earnings, D/E, and EP metrics for all programs, both GE and non-GE. The GE Information Rates Data included records for non-GE programs and the number of completers from IPEDS, but did not include any performance information for non-GE programs.
- 2022 PPD defines the universe of programs and measures the number of completers based on ED administrative data held in NSLDS. The GE Information Rates Data used information derived from IPEDS to create the universe of programs so the set of programs in the two data sets differs slightly.
- 2022 PPD includes information about student-level characteristics and borrower outcomes that were not previously included in the GE Information Rates Data.

The variables included in the 2022 PPD are constructed using different methodology than that described in the proposed regulations in several ways. Briefly, these differences are:

- 2022 PPD uses 4-digit CIP code to define programs, but the proposed rule uses 6-digit CIP codes. The RIA has a discussion of the differences in coverage using 4-digit vs. 6-digit CIP codes to define programs.
- The proposed rule would cap individual students' total loan debt at net direct expenses (tuition, fees, books, and supplies less institutional grant aid provided to the student) and also include non-federal loans when computing individual debt; the 2022 PPD does not adjust debt for either of these.
- Under the proposed rule, program median debt and earnings would be computed using the same two-year cohort of completers, though the 2022 PPD used slightly different cohorts to compute these metrics.
- Medical and dental professional programs with post-graduation residency requirements are not evaluated in the 2022 PPD, but would be evaluated under the proposed rule, albeit with a longer time horizon for earnings measurement as described in the NPRM.
- The proposed rule would use a four-year completer cohort to compute metrics if the two-year cohort did not meet the n-size requirement of 30 completers. The 2022 PPD did not compute four-year cohort outcomes.

For a more in-depth discussion of the ways that the 2022 PPD data differs from that provided at negotiated rulemaking or from the proposed methodology, see the "Data Used in this RIA" section of the proposed rule.

Privacy Protocols for Release of Public 2022 PPD

To protect student privacy, the Department applied certain protocols to the 2022 PPD to construct the Public 2022 PPD; this public dataset thus differs somewhat from the 2022 PPD analyzed in the Regulatory Impact Analysis. These protocols are standard practice for the Department when releasing

data derived from administrative student records held by the Department. The privacy protocols include four steps applied to variables that are derived from Department administrative data:

- For most variables included in the 2022 PPD, we identify the number of students used to construct each variable. Whenever this count is less than 30, the value of the variable is suppressed by setting it equal to missing in the Public 2022 PPD.
- For any student count variables included in the 2022 PPD (e.g., count_AY1617, earn_count_ne_3yr), the value is set to 25 if the count was between 20 and 29 and set to 15 if between 10 and 19. Counts less than 10 are suppressed. The count variable avg_t4_enr_AY1617 averages enrollment across two years, so it should be doubled to determine the relevant number of students for privacy purposes. Thus, the value of this variable in the Public 2022 PPD is set to 13 if avg_t4_enr_AY1617 was between 10 and 14, and to 8 if between 5 and 9; counts less than 5 are suppressed in the Public 2022 PPD.
- Variables that correspond to shares (e.g., the share of students ever receiving Pell) require additional protocols on top of those described above, so that counts of small groups cannot be inferred from the shares. For these variables, we compute the number of individuals in each group implied by the share (and 100% minus the share) when the share is multiplied by the total count. When this implied count is between 20 and 29, we compute a new share equal to 25 divided by the total count. If the implied count is between 10 and 19, we compute a new share equal to 15 divided by the total count. If the implied count is less than 10, the value for the share variable is suppressed entirely.
- Any variables that are suppressed for more than half the records in the 2022 PPD after applying the above steps are not included in the Public 2022 PPD. Analysis using these highly suppressed variables could provide a misleading view of the impact of the regulation given their selective coverage.

The accompanying codebook identifies variables to which the above privacy protocols have been applied and for which some values have been suppressed.