Report and Suggestions from College Scorecard Technical Review Panel 2: Program-Level Measures and Institutional Comparisons

The Technical Review Panel met to discuss issues related to improvements to the College Scorecard. This summary provides feedback from the group on how measures at the program level could be calculated and added to the Scorecard as well as how comparative information could be included on the current Scorecard website for institution-level metrics. Comments from interested parties are due to Janice Kelly-Reid, Project Director at RTI International, at ScorecardTRPcomment@rti.org by June 20, 2019.

RTI International (RTI) is under contract to the U.S. Department of Education (Department) to form a Technical Review Panel (TRP) and conduct meetings to solicit expert discussion and suggestions on a range of topics related to the College Scorecard. The TRP is designed to provide input to RTI on data, measures, and tools to help inform ongoing improvements to the College Scorecard. The TRP does not advise or report to the Department.

On April 11 and 12, 2019, RTI International convened a meeting of the TRP in Washington, DC. The purpose of this TRP was to solicit expert advice from other organizations (i.e., federal agencies, postsecondary institutions, research organizations, and other stakeholders) informing the methodology for calculating some of the program-level[1] data metrics that will eventually be presented in the College Scorecard related to the recent executive order on Improving Free Inquiry, Transparency, and Accountability at Colleges and Universities. In addition, expert panelists were asked to provide advice on the presentation of key metrics on the College Scorecard consumer website. The panel consisted of 42 individuals representing institutions, researchers and other data users, higher education associations and think tanks, the federal government, and other experts, reflecting a wide range of expertise and a diverse array of perspectives on this topic.

Background

The College Scorecard is an open data, consumer information initiative that includes these components:

- The College Scorecard data site with downloadable data files and the Application Program Interface (API) which allows developers to integrate updated data into their own applications.
- The College Scorecard consumer tool, built from the API, helping students, parents, and counselors make informed college choices.

The College Scorecard data site enables researchers, policymakers, and others to conduct in-depth analysis, and enables institutions to use the data to benchmark against peer institutions. The entire dataset includes over 2,000 data elements spanning nearly 20 years of information from approximately 7,000 institutions and covers multiple federal sources, including the Integrated Postsecondary Education Data System (IPEDS), the National Student Loan Data System (NSLDS), and administrative earnings data from tax records maintained by the Department of the Treasury.

[1] For the purpose of the TRP discussions, a program is defined as an award level (e.g., certificate, associate's degree, bachelor's degree, graduate-level degrees) and the instructional program as defined by the Classification of Instructional Programs (CIP).

The College Scorecard consumer tool highlights a subset of College Scorecard data elements and enables users to search for and evaluate options by presenting comparable and accessible data on postsecondary institutions and their performance on measures of access, affordability, and student outcomes. For each institution, the consumer tool includes data elements describing general characteristics (e.g. location, size, etc.), costs and debt, graduation and retention, earnings after attendance, student characteristics, admissions requirements, and academic programs available.

Roughly 3.8 million unique users have visited the College Scorecard consumer site since September 2015. There have been approximately 130,000 downloads of the College Scorecard data files. Nearly 5,000 developers have used the Scorecard API.

Why Program-Level Outcomes Information Matters

Before discussing program-level measures, panelists discussed why consumers might find program-level information useful in addition to institution-level data already available on the College Scorecard. The panelists reviewed analysis of earnings outcomes of college graduates. Key findings indicate selecting both an institution AND a field of study are important decisions for prospective students and significantly affect long-term earnings. For example, research from the Georgetown University Center on Education and the Workforce shows that, nationally, 10 years after graduation, those with bachelor's degrees earn a median of $62,000. Depending on a student's program of study, however, earnings can vary greatly—from a median of $42,000 for a graduate in education to $85,000 for a graduate in architecture or engineering. Even within program categories, there is often substantial variation. For example, median earnings for graduates in early childhood education are $40,000, while median earnings for graduates in physical and health education are $52,000. In engineering, median earnings vary from $68,000 for architectural engineering to about $140,000 for petroleum engineering. Analysis also showed there was substantial variation in the earnings outcomes of students within an institution, based on their program of study.

The panelists generally agreed that adding program-level outcomes information will enable students to compare potential outcomes among programs at a single institution as well as across institutions. Panelists noted that this information will also be useful to institutions themselves, which now often depend on third parties to obtain this information about their graduates, sometimes at considerable cost.

Program-Level Reporting in Federal Student Aid Data System

In order to understand the underlying data used to calculate program-level metrics, panelists reviewed the details of how institutions report program-level data to the Department. Until recently, the Department did not have the necessary information to calculate information on the College Scorecard at the program level. However, beginning in 2014, the Department began collecting program-level information on all federal financial aid recipients in the National Student Loan Data System (NSLDS), and the data needed to calculate outcome measures based on program level are now available.

Status of Program-Level Reporting to NSLDS

In order to examine data quality implications of deriving program-level data metrics from NSLDS, panelists reviewed analyses comparing program-level enrollment data in NSLDS with other sources of

data. Panelists reviewed analyses comparing program-level enrollment records reported in NSLDS with institution-level enrollment records in NSLDS. Institution-level graduation records serve as a reliable benchmark because these data are used operationally to determine when federal loans go into repayment. Moreover, each institution-level graduation record should have a corresponding program-level graduation record describing which program a student completed. To estimate the completeness of program-level reporting, panelists reviewed analyses describing the percentage of institution-level graduation records with a corresponding program-level graduation record. Analysis suggests that program-level graduation reporting is fairly complete in recent years but less complete in the earliest year of data collection. For example, in award years 2016-17 and 2017-18, the typical institution reported a corresponding program-level graduation record for more than 95 percent of institution-level graduation records. However, the match rate of the typical institution was less than 80 percent in award year 2014-15.

Panelists reviewed analysis comparing program-level enrollment data in NSLDS with data from the National Center for Education Statistics (NCES) postsecondary surveys: National Postsecondary Student Aid Study (NPSAS), Beginning Postsecondary Students Longitudinal Study, and Baccalaureates and Beyond Longitudinal Study, collectively referred to as the sample surveys because only a portion of students at selected Title IV institutions are statistically chosen to be surveyed. The sample surveys' data are used to generate a nationally representative portrait of students enrolled in postsecondary education and track their experiences and outcomes. Comparisons with the sample surveys suggest that NSLDS program-level data correspond reasonably well, in terms of a match rate, but coverage varies by aid type and credential level. The analysis suggests program-level enrollment records for non-borrowers (e.g., Pell-only students) and students in certificate programs are less complete.

Panelists also examined analysis comparing program-level reporting in NSLDS to completions data reported at the institution level to IPEDS. Although there are some programs in the NSLDS program-level data file that did not match the completions data, the analysis suggests high correlation between the data sources, both at the 4-digit CIP level and the 6-digit CIP level. Panelists indicated that individual institutions may report awards conferred differently to NSLDS from how they report awards conferred to IPEDS, which may explain the outliers.

Based on the various analyses, panelists felt most confident in NSLDS's reporting for bachelor's degree students and less comfortable with its reporting of certificate programs, at both the undergraduate and graduate levels. Generally, the panelists suggested if subsets of programs remained problematic, College Scorecard should consider phasing in the program-level information being added to the College Scorecard. For example, starting with the programs for which the coverage seems to be best (bachelor's and associate's degrees).

Discussion: Defining Programs (CIP-4 vs. CIP-6)

In order to consider the most appropriate way to define a program, panelists examined the landscape of programs and institutions. The Classification of Instructional Programs, or CIP, provides a taxonomic scheme that supports the accurate tracking and reporting of fields of study and program completions activity. Three nested levels of CIP codes are available: 2-digit, 4-digit, and 6-digit. The 2-digit code is a broad category, while the 4-digit and 6-digit codes get more specific. Among programs with at least one award conferred in NSLDS, roughly 15 percent of 4-digit CIP code programs contain more than one 6-

digit program, and nearly 70 percent of awards were in 4-digit CIP code programs with only one reported 6-digit CIP code program.

An important consideration when determining at what CIP level to provide program-level information is cell size. Data elements are suppressed for groups few students. Panelists reviewed analysis comparing the percentage of programs with publishable data on loan debt at the 4- and 6- digit CIP level. For the purpose of publishing mean loan debt at the program level on borrowers, for example, approximately 51 percent of programs would be suppressed when reporting at the 4-digit level and 56 percent would be suppressed when reporting at the 6-digit level. However, the percentage of awards in the unsuppressed programs is very high: 87 percent at the 4-digit level and 85 percent at the 6-digit level.

Another consideration when deciding at what CIP level to report program information is how states are already reporting program-level outcomes. Panelists reviewed the findings of an environmental scan of methodology for program-level metrics. The analysis revealed a diversity of methodologies among states with no consistent standard on the appropriate unit of analysis.

Additionally, panelists considered how 6-digit and 4-digit CIP codes would be interpreted by prospective students comparing CIP names to program titles in course catalogs. Analysis of college search user behavior suggests exploration in search happens at the 2-, 4-, and 6-digit CIP levels. In addition, data at the 6-digit CIP code level can be more difficult to interpret because there may not be a clear mapping between the 6-digit CIP code and the course catalog program names.

After much discussion, the panelists generally agreed the 4-digit CIP code was a good balance between providing information on a granular level while maintaining student privacy. However, some panelists were concerned that since multiple programs could be captured under the same 4-digit CIP code, calling it “program-level information” could be misleading to students and parents. Therefore, panelists suggested that the College Scorecard provide information at the 4-digit CIP code level but use a term other than program-level information. One suggestion is that it be referred to as “field of study” instead of “program.”

There was also discussion about whether measures should be calculated at additional levels of aggregation, for example, for undergraduate versus graduate or by credential level. Panelists did not agree whether College Scorecard should do so.

Discussion: Program-Level Measures—Earnings

College Scorecard plans to publish post-enrollment earnings data by field of study for cohorts of students who graduate in a two-year period starting in the fall of 2019. In order to consider how program-level median earnings of graduates should be calculated, panelists reviewed an overview of the draft cohort requirements using graduation data by CIP code and credential level reported to NSLDS. The draft requirements exclude students who were enrolled in school during the measurement year, had military service during the measurement year, and students who died or had a disability discharge prior to the end of the earnings measurement year. The panelists provided feedback on the methodology and agreed that non-earners should not be included in the median earnings calculation, but also agreed that some information regarding the percentage of graduates who are non-earners could be provided for context. For example, the College Scorecard could report the percentage of graduates who are subsequently enrolled in an institution and the percentage who are active in the military and thus are not included in the earnings figures.

In order to determine the best time to measure earnings, panelists considered analysis and lessons learned from states regarding factors to consider in choosing when to measure outcomes. . Panelists suggested College Scorecard provide both short-term and long-term earnings data. However, NSLDS does not yet have enough data on historical earnings prior to 2014. The panelists agreed that the College Scorecard should consistently report first-year earnings post-completion each year for consecutive cohorts and should provide additional years of earnings over time as the data become available.

Discussion: Program-Level Measures—Debt

Panelists reviewed the draft methodology for calculating median and mean cumulative student loan debt, based on NSLDS graduation events by CIP code and credential level. This measure represents the sum of all federal loan disbursements made through the Direct, FFEL (Federal Family Education Loan), and Graduate PLUS loan programs for graduates. Loans that were disbursed by the different institution would not be included, nor would Parent PLUS loans, Perkins Loans, and loans that were disbursed at a different academic level (i.e., graduate versus undergraduate). Panelists discussed whether non-borrowers should be included in the debt measures. An example of a student in NSLDS who is not in the borrower-only cohort is a student who received a Pell Grant but not a federal loan. Panelists reviewed analysis of privacy implications and the impact on including non-borrowers into the calculations in terms of the amount of publishable data. As with other NSLDS calculations, the data are only for Title IV aid recipients. Given this limitation of the data, panelists suggested that non-borrowers should not be a part of the debt calculations. However, they agreed that publishing the count or percentage of students who are borrowers would provide important context for consumers looking at the data. For example, IPEDS provides the percentage of undergraduates at the institution-level who receive federal loans. Panelists cautioned against combining NSLDS data with IPEDS completion survey data to calculate frequency of borrowing at the program level because institutions may not be reporting consistently to both data systems. However, panelists suggested College Scorecard consider using the two data systems to approximate ranges of students who borrow and encourage institutions improve consistency in reporting over time so that calculations could be more precise in the future.

Discussion: Program-Level Measures—Repayment Rate

Panelists discussed possible definitions of a repayment rate, including a borrower-based rate (percentage of borrowers without a prior default lowering the balance by $1 within a set period of time), a dollar-based rate (percentage of program portfolio paid down between within a set period of time), and the median percentage paid down over a set number of years among borrowers in the cohort.

Panelists brought up the fact that borrowers can be in “good standing” on their loans but not be paying down the principal because they are in an income-based repayment program and their income is not yet high enough to need to pay on the loan.. Panelists discussed ways to characterize such borrowers who for some reason do not yet have to pay on their loans.

Most of the discussion centered around additional information that could be provided to show when borrowers are in good standing regardless of whether they are paying down their principal. However, there was no general agreement as to what “good standing” means. Some panelists suggested that College Scorecard provide the percentage of borrowers who are in income-based repayment plans for

context. Some panelists suggested that the repayment rate in College Scorecard could display the percentage of students in several different categories, including default, delinquency, negative amortization, and positive amortization. The Department noted that calculating frequency of income-based repayment plans and delinquency statuses would only be possible for Direct Loans, not FFEL loans.

Discussion: Comparison Points on the Consumer Tool—How Should Current Key Institution-Level Measures Be Presented?

The final session of the TRP meeting was dedicated to discussing how to add information to the College Scorecard that students could use to compare institutions. The College Scorecard used to provide national median comparison lines for key institutional measures. In 2018, these comparison lines were removed to minimize potentially misleading comparisons among different types of institutions with different key characteristics. The Panelists reviewed analysis describing the distribution of data values across different groupings of institutions and discussed alternative institutional comparison points to include on the College Scorecard to provide more context to consumers.

Panelists considered providing comparison points for institutions based on many factors: the main credential offered by an institution and its admissions policy; location and related cost of living; and College Scorecard users' search terms. While most panelists preferred including a comparison, the group did not agree on what would be a good comparison point.

Next Steps

Once the TRP summary comment period has closed, RTI will review the comments and outline recommendations for the Department based on the outcome of the TRP meeting and subsequent public comment period. The Department will review the recommendations to determine changes for future versions of the College Scorecard.

Comments

RTI is committed to improving the quality and usefulness of the College Scorecard. We encourage interested parties to send any comments or concerns about this topic to Janice Kelly-Reid, Project Director at RTI International, at ScorecardTRPcomment@rti.org by June 20, 2019.

.