

Towards a Framework for Accountability for Federal Financial Assistance Programs in Postsecondary Education

Jordan Matsudaira and Lesley J. Turner¹

Appendix B: Simulation Methodology

B.1 Loan repayment rate metric

The data provided by the Department of Education (ED) includes cohort loan balances at repayment and one, two, and three years later, as well as cohort loan balances for students who are either in school, have a military deferment, or are not required to repay their loans for other reasons (e.g., death or disability) at repayment and one, two, and three years later. Repayment rates are calculated for programs defined based on 2-digit CIP code by credential level by institution. Post-baccalaureate certificate and graduate certificate programs are combined into a single category.

B.2 Net earnings premium metric

Ideally, the measure of cohort earnings used in the construction of the net earnings premium would be median earnings for all Title IV exiters in a given cohort, measured 3 years after they left the program (irrespective of completion status). However, this data is not currently available. To estimate this quantity, we start with the College Scorecard's program-level earnings data. These data are reported by 4-digit CIP code level for 8 different types of credentials: undergraduate certificates, associate degrees, bachelor's degrees, post-baccalaureate certificates, graduate certificates, master's degrees, doctoral degrees, and first professional degrees. Due to the small number of post-baccalaureate and graduate certificate programs, we combine these categories into a single "graduate certificate" category.

We first convert the College Scorecard's measure of earnings one year after exit to estimated earnings 3 years after exit. To make this adjustment, we use estimates of annual earnings growth rates by field and credential level from the American Community Survey (ACS). Specifically, we estimate the annual growth rate of individuals with Bachelor's degrees between the ages of 25 and 30 based on the CIP code of their Bachelor's degree, and then adjust these growth rates for non-BA programs based on

¹ Matsudaira: jm4763@tc.columbia.edu; Turner: lesley.j.turner@vanderbilt.edu. The paper that these appendices accompanies can be found at: <https://www.brookings.edu/research/towards-a-framework-for-accountability-for-federal-financial-assistance-programs-in-postsecondary-education/>.

differences in average earnings growth rates across degree levels (e.g., if the estimated growth rate of bachelor's degree recipient earnings are 1 percent greater than that of associate degree holders, we would assume that the earnings growth for all associate degree programs would be the same as the CIP-specific bachelor's degree growth rate minus 1 percent).

Our ideal earnings metric would include both graduates and non-completers, but earnings reported in the College Scorecard are for program completers only. To convert earnings for completers to earnings for both completers and non-completers, we use data on the average earnings of borrowers by completion status measured 4 or 6 years after students begin their studies.² We use these data to estimate the 6-year completion rate of Title IV borrowers and the ratio of non-completer earnings to completer earnings. We apply these institution-level ratios to year-three earnings calculated in our first step for all programs within the same institution. When institution level data are missing, we impute them based on sector level averages. This provides an estimate of median earnings for all exiters including non-completers, measured 3 years after exit.

The second component of the net earnings premium metric is median out-of-pocket costs. To estimate out-of-pocket costs for undergraduate programs, we first use College Scorecard data on the share of all undergraduates that reenroll, graduate, and drop-out in their second, third, fourth, fifth, seventh, and ninth year since entry. We use IPEDS data to calculate each institution's net tuition and fees per student. Per student net tuition is then multiplied by the share of students who remain enrolled 2 through 9 years after entry and these quantities are summed to measure school-wide total out-of-pocket tuition and fee costs per student. Next, we calculate the ratio of median debt to estimated out-of-pocket costs per student. Finally, this ratio is applied to program-level median debt to produce an estimate of undergraduate program-level total expected out-of-pocket costs. For graduate programs, we use median debt as a proxy for out-of-pocket costs. Total out-of-pocket costs are converted into an annual payment by amortized the total amount over a 20-year period for undergraduate programs and over a 25-year period for graduate programs.

The third component of the net earnings premium metric is counterfactual earnings. For undergraduate programs, we use median annual earnings of high school graduates and GED holders aged 25 to 34 who live in the same state as the institution. Counterfactual earnings for graduate

² These data were produced by Matsudaira using the deidentified data underlying the release of the original College Scorecard in 2015. These data report institution level information based on students who began their studies in 2006, with earnings measured in 2010 or 2012.

programs are set equal to median earnings of individuals aged 25 to 34 in the same state who have a bachelor's degree in the same broad category as the program.³ These categories are:

- Arts and Humanities
 - ACS categories: Architecture; Area, Ethnic, and Civilization Studies; Linguistics and Foreign Languages; English Language, Literature, and Composition; Liberal Arts and Humanities; Interdisciplinary and Multi-Disciplinary; Philosophy and Religious Studies; Theology and Religious Vocations; Fine Arts; History.
 - CIP codes: 4, 5, 16, 23, 24, 30, 38, 39, 50, 54.
- Education and Public Service
 - ACS categories: Education Administration and Teaching; Library Science; Psychology; Criminal Justice and Fire Protection; Public Affairs, Policy, and Social Work.
 - CIP codes: 13, 25, 42, 43, 44.
- Agriculture, Consumer Services, and Trades
 - ACS categories: Agriculture; Environmental and Natural Resources; Communications; Communication Technologies; Cosmetology Services and Culinary Arts; Family and Consumer Sciences; Physical Fitness, Parks, Recreation; Construction Services; Electrical and Mechanical Repairs and Technologies; Transportation Sciences and Technologies.
 - CIP codes: 1, 3, 9, 10, 12, 19, 31, 46, 47, 48, 49.
- Business and Social Sciences
 - ACS categories: Law; Social Sciences; Business.
 - CIP codes: 22, 45, 52.
- STEM and Health
 - ACS categories: Computer and Information Sciences; Engineering; Engineering Technologies; Biology and Life Sciences; Mathematics and Statistics; Military Technologies; Physical Sciences; Nuclear, Industrial Radiology, and Biol ; Medical and Health Sciences and Services.
 - CIP codes: 11, 14, 15, 26, 27, 28, 29, 40, 41, 51, 60.

³ Appendix Figures A5, A6, and A7 display median earnings for high school graduates by state, median earnings of bachelor's degree holders by state, and median earnings of bachelor's degree holders by state and broad field, respectively.

B.3 Number of Title IV exiters

We estimate the number of Title IV exiters using tabulations from ED on the number of borrowers entering repayment at the undergraduate and graduate-level in 2016 and IPEDS data on the number of program completers for the same year.⁴ To do this conversion, we need information on the share of all completers that are TIV recipients and the share of all exiters that are completers.

From the 2016 NPSAS, we calculate the share of Title IV recipients that did not receive federal loans by institutional control and highest degree offered (9 categories). We apply this ratio to the number of undergraduate borrowers entering repayment to get an estimate of the total number of Title IV recipients leaving the institution in a given year. At the graduate level, we assume that all Title IV exiters were borrowers and thus, the count of graduate borrowers entering repayment provides the count of Title IV recipients. For schools with both undergraduate and graduate students, we sum these two figures for an estimate of the total number of Title IV students who left college in 2016. To approximate the number of Title IV exiters at the program-level, we take the ratio of the number of school-level completers to the estimated number of school-level Title IV exiters and divide total program-level completers by this ratio. We assume that all programs that report the number of completions and median earnings have at least 10 Title IV exiters. We assume that programs for which both median earnings and the number of completers is suppressed have 5 Title IV exiters.

The estimated number of exiters at the 4-digit CIP code level are summed to calculate the number of exiters at the 2-digit CIP level. When both enrollment and earnings are missing for all 4-digit CIPs within a 2-digit CIP that has a nonmissing loan repayment rate, we impute enrollment based on a Poisson regression of Title IV exiters on the institutional predominant degree, 2-digit CIP code, program credential level interacted with loan volume at repayment entry, number of program completers, and number of estimated Title IV exiters at the institution level, separately by control. This regression is run on all program-level observations that have program-level Title IV exiters and then the values of the coefficients are used to predict program-level Title IV exiters for programs where this is missing.

Approximately 9 percent of all (estimated) TIV exiters attend programs that have neither a repayment rate nor a measure of net earnings.⁵ Around 10 percent of this group is enrolled in an institution that does not participate in federal loan programs, which means that these exiters might be

⁴ The IPEDS data are publicly available while data on the number of borrowers entering repayment by level was provided to the Senate HELP Committee by ED and shared with the authors upon request.

⁵ Note that this is different than earlier estimates of the number of *completers* in programs that did not have earnings and/or loan repayment rate metrics.

covered by the proposed accountability metrics if earnings data were reported at the 2-digit rather than at the 4-digit CIP level. Note that 21 percent of Title IV exiters have a repayment rate but no earnings data, likely representing enrollment in programs that have enough exiters to report at the 2-digit but not at the 4-digit level. Finally, 0.4 percent of exiters are in programs that have earnings but not loan repayment rate (but 70 percent of these programs are in schools that do not participate in loan programs). Table A1 displays the breakdown of programs by credential level that are missing both metrics at the 2-digit CIP-level.⁶

We construct weights for the set of programs that have a valid net earnings premium, a valid loan repayment rate, or both metrics that will make the subset of programs for which we have either or both metrics representative of enrollment across all programs. These weights are generated via raking, where the distribution of enrollment in each subset of programs described above will match the distribution of total enrollment by field (2-digit CIP code), institutional control, institutional predominant degree, and credential level when the weights are applied.⁷

B.4 Sample

We exclude programs in schools that are closed from our main analysis sample. There are 571 institutions that have programs for which the loan repayment rate or net earnings premium can be measured but are no longer operating. These schools were predominantly for-profits, which make up 474 of the closed institutions. The remaining institutions include 92 nonprofit institutions and 5 certificate-granting public institutions.

Many borrowers who attended a subset of the closed for-profit institutions received partial or full federal student loan forgiveness either through automatic closed school discharges (ACSD) or the earlier borrower defense to repayment (BDR) regulations.⁸ These institutions include schools owned by Corinthian Colleges (Everest, Heald, Art Institutes, Wyotech), and locations of Argosy University, ITT Technical Institute, and many other large for-profit chains. In many cases, student loan discharges occurred within the three years after a borrower had entered repayment and thus, programs in these institutions could appear to have artificially low loan repayment rates.

⁶ An alternative approach to aggregating data to the 2-digit CIP level would be to aggregate exiter cohorts. The trade-off to this approach is that more time would elapse before a program with poor student outcomes faced accountability consequences, resulting in more students suffering the consequences of poor performing schools.

⁷ We use the Stata command `ipfraking` to estimate weights.

⁸ See <https://studentaid.gov/sites/default/files/fsawg/datacenter/library/ACSD.xls> and <https://studentaid.gov/data-center/student/loan-forgiveness/borrower-defense-data> for additional details and data.

Appendix C: Frequently Asked Questions and Answers

1. *What about student demographics? Should we be concerned that sanctions based on the proposed accountability metrics will cause disproportionate harm to institutions and programs that serve large shares of low-income students and students of color simply because of student demographics rather than their performance? Should we worry about incentives for institutions to deny admission to any groups of students?*

Answer: We argue that a “lower bar” approach for determining Title IV federal student aid eligibility largely negates this concern. The proposed thresholds are intended to represent the bare minimum required to be reasonably sure that most students are not harmed by attending a program. If this is indeed true, we do not want to hold programs that serve different populations to different standards (or, in other words, accept worse outcomes as acceptable for students from disadvantaged backgrounds and those who are underrepresented minorities). If we think that the proposed thresholds are too high, that is an argument for moving the thresholds for all programs rather than just for schools serving specific students.

As discussed in Section 7.2, demographics are not destiny. Among schools that are in the top quartile of both Pell Grant receipt and the share of the undergraduate student body that are under-represented minorities, 13 percent of programs have net earnings in the top quartile of the NEP metric for the program’s credential level and 12 percent of programs have loan repayment rates in the top quartile.

The adverse incentives for institutions to change their admissions standards and/or reject particular applicants would likely be a bigger issue with higher thresholds. In practice, almost all of the undergraduate programs that are predicted to fail are situated within open-access schools that do not have the ability (either due to their mission or because of finances) to deny admission to any group of students. Additionally, as we discuss in Section 8, most failing programs are in schools with multiple non-failing programs (see also Table 10). Although we cannot test the extent to which within-school differences in program performance correlate with differences in the demographics of students enrolled in these programs, the degree of heterogeneity in outcomes suggests that in many cases, student characteristics may be less important than the field of study and credential that is pursued.

2. *Should program exiters with \$0 earnings be excluded? If a large portion of students are unemployed (in the labor force and looking, but not finding it) then shouldn’t this be taken into account when evaluating performance?*

Answer: Ideally, exiters that are unemployed but remain in the labor force would be included in the calculation of median program net earnings. However, data limitations make this difficult. Earnings data would most likely come from the IRS, which does not have sufficient information to determine whether an individual with \$0 earned income is unemployed or out of the labor force.

Theoretically, it might be possible to get information on who received unemployment insurance benefits from the IRS through Form 1099-G. But this would still exclude those who didn’t qualify for unemployment insurance (i.e., those who were fired for cause) and the long-term unemployed. While only measuring median earnings for exiters with any annual earnings may let some programs off the

hook, erring on the side of making the NEP more conservative is consistent with the lower bar approach we have taken in constructing the proposed accountability metrics.

3. *What about interactions with income-driven student loan repayment (IDR) plans? If more students enter IDR or if students are automatically enrolled in IDR, as has been proposed by some policy makers, then what does it mean for repayment? Should the loan repayment rate be adjusted for IDR participation in some way?*

Answer: As IDR participation increases, the repayment rate metric becomes a better measure of student well-being.⁹ The share of students that have an “on-time repayment” status (i.e., those not in default/delinquent/forbearance) is not necessarily a measure of the share of students who are better off from having borrowed to attend the program. Theoretically, IDR participants will have their remaining debt forgiven at the end of the repayment period, so someone who never makes payments on their loans because their earnings are persistently low is not necessarily worse off than someone who never borrowed to attend the program. But this argument abstracts from any psychic costs of holding student debt for such a long period and real effects on access to other sources of credit through effects on a borrower’s debt-to-income ratio.

While we do not observe program-level IDR participation, we are able to separately calculate LRRs and IDR participation rates for a school’s undergraduate and graduate borrowers. Average undergraduate IDR participation rates are quite similar for schools with a positive and negative undergraduate LRR. Specifically, 21% of undergraduate borrowers (30% of undergraduate loan dollars) from schools with a positive undergraduate LRR were in an IDR plan while 20% of undergraduate borrowers (29% of undergraduate loan dollars) from schools with a negative undergraduate LRR were in IDR. Differences in IDR take-up by a school’s graduate LRR are present, however: in schools with a passing graduate LRR, 21% of graduate borrowers (30% of graduate loan dollars) are in IDR, while in schools with a negative graduate LRR, 25% of graduate borrowers (37% of graduate loan dollars) were in one of these repayment plans.

4. *Why use a dollar-based based repayment rate instead of a borrower-based repayment rate? The dollar-based metric takes the focus away from the share of students who are better/worse off and allows students who are repaying well to compensate for those who are doing poorly.*

Answer: This is an important concern given our first principal of design for accountability metrics: that such measures be linked to outcomes that are clearly and unambiguously good for students. At first glance, a borrower-based repayment metric does a better job of meeting this goal than a dollar-based metric. However, there are circumstances under which a borrower-based metric would not fully capture the extent to which a cohort of borrowers are better or worse off. As an example, take a cohort that is comprised of 6 borrowers with \$5000 in debt and 5 borrowers with \$80,000, and assume that the

⁹ However, under the three most popular IDR options, the federal government covers a portion of unpaid interest when the borrower’s required payment is not sufficient to cover both principal and interest in the initial three years of repayment. Specifically, 100% of subsidized loan interest is covered for borrowers in REPAYE, PAYE, and IBR. Borrowers in REPAYE will also have half of the unpaid interest on unsubsidized loans for this period. However, borrowers must recertify their income each year to continue to participate in IDR and those who fail to do so will have any unpaid interest capitalized into their loan principal. See <https://studentaid.gov/manage-loans/repayment/plans/income-driven/questions> for additional details.

threshold for the borrower-based metric is 50% (i.e., that at least half of the individuals in the cohort have reduced their debt balance by \$1). If the low-balance group reduced their balances by a dollar but the high balance group saw their balances all increase by 10% (\$8000), the cohort as a whole would not pass the dollar-based metric but would pass the borrower-based metric. In contrast, suppose the borrowers with \$1000 in debt saw their balances increase by \$1 and the high balance borrowers paid down 10% of their debt. This cohort would fail the borrower-based metric but pass the dollar-based metric. While this is perhaps an extreme example, it illustrates the ability of the dollar-based metric to require a relatively larger payments by higher balance borrowers before a school is classified as passing.

Additionally, while the threshold for the dollar-based repayment rate is straightforward – whether the cohort has made any progress paying down their student loans in aggregate – there is not an obvious threshold for the borrower-based metric. One option would be to require that at least half of a cohort of borrowers to reduce their debt by at least \$1. While this seems like a relatively low bar, at the school-level, a much larger share of institutions would “fail” based on this metric than under the dollar-based repayment rate (51% versus 27% of all schools).¹⁰

Finally, a dollar-based metric provides incentives for schools to reduce the amount students need to borrow (which is different than reducing the share of students who borrow). Thus, there are at least indirect incentives to lower and/or control costs.

5. *Should the repayment rate metric be adjusted for the percentage of students who borrow?*

Answer: There are both beneficial and harmful ways to reduce borrowing. The inclusion of net out-of-pocket direct costs in the earnings metric provides incentives for programs to reduce costs, which should in turn reduce the need to borrow through a beneficial channel. Our concern is that including an adjustment for the percentage of students who borrow into the repayment rate metric would provide incentives for schools to reduce borrowing through methods that have been shown to harm students, such as opting out of federal loan programs or not packaging loans (Dunlop 2013, Weiderspan 2016, Marx and Turner 2019, Barr et al. 2020).

6. *What about bachelor’s degree seeking students? Many do not declare a major until their second or third year. How would drop-outs be classified with respect to program if they drop-out before declaring a major?*

Answer: We briefly discuss potential solutions to this concern in Section 9.2. One option would be to only calculate program-level accountability metrics for bachelor’s degree-seeking students who have accumulated more than a threshold number of credits (e.g., 1 to 2 semesters of full-time equivalent enrollment). This approach would work best if paired with an institution-level backstop to avoid adverse incentives for schools offering bachelor’s degree programs to encourage students perceived to be risky to drop-out before this threshold is met.

7. *How should the metrics account for transfers? E.g., What if debt that was taken on in a for-profit associate degree program and then student transferred to a bachelor’s degree program in a nonprofit*

¹⁰ The threshold in borrower-based loan repayment that would lead to similar rates of institutional failure as the measure based on whether cohort balances have fallen by \$1 is 37%.

institution? Will the debt taken out for the associate program be allocated to the bachelor's degree program or institution?

Answer: Debt will be attributed to the program that the student was enrolled in when they borrowed. There will be some complications for a programmatic loan repayment metric (e.g., double majors, students that switch programs mid-semester) but for students who transfer between programs at different institutions, this should be straightforward, as federal loans are originated by institutions. The attribution of borrowers to exit cohorts for the purpose of calculating the LRR would be the same as how undergraduate borrowers are currently attributed to repayment cohorts in CDR calculations. For instance, bachelor's degree recipients who immediately enter a graduate program would be counted in both the bachelor's degree program exit cohort and graduate program exit cohort when the student leaves the graduate program and enters repayment.

8. How are online students handled in each metric?

Answer: For fully online programs, we propose using the national median for counterfactual earnings in the NEP metric. For partially online and hybrid programs, the national median could also be used if the percentage of credit hours/instructional hours exceeds some threshold (e.g., 50%). There is no need to adjust the loan repayment rate metric.

9. Should you adjust for shorter vs. longer certificate programs? Or, more generally, use finer degree categories such as BA vs BS, MA vs MS, etc.?

Answer: This will ultimately depend on the extent to which this data is currently being collected (or can be collected in the future). There are two additional considerations. First, finer delineations of programs will result in smaller cohort sizes. Some programs that would have been sufficiently large to have metrics calculated with the broader degree categories will become too small. We have proposed the solution of pooling together several cohorts when this occurs, but the pooling process will necessarily delay when sanctions are imposed on low-performing programs.

10. What about programs that lead to occupations where tips make up a large share of earnings (and tips are underreported)?

Answer: The under-reporting of tipped income is one of the main reasons for using a metric based on loan repayment as a second measure of economic well-being. Indeed, a large share of sub-baccalaureate service programs (a 2-digit CIP code that includes cosmetology and hospitality programs) have a negative net earnings premium and a positive loan repayment rate (Figures 2 and 3). Specifically, of the undergraduate certificate programs in this field, 8 percent have both a positive net earnings premium (NEP) and loan repayment rate (LRR), 37 percent have a negative NEP and LRR, less than 2 percent have a positive NEP and negative LRR, and the largest share – 53 percent – have a negative NEP and positive LRR.

11. How noisy are these measures? How much does performance vary year to year?

Answer: Using a supplemental data set, we construct a school-level measure that is a close approximation of the loan repayment rate (LRR) in our proposal for the 2009 through 2011 cohorts.¹¹ We use this data to test whether the likelihood that a school “fails” this LRR metric at the school-level is highly correlated across cohorts. Among public institutions, 84 percent pass in all 3 years, an additional 9.5 percent pass in 2 out of 3 years, and 2.6 fail percent in all 3 years, and the remaining 4 percent pass in one out of 3 years. Among nonprofit institutions, things look quite similar: 84 percent pass in all 3 years, an additional 6.7 percent pass in 2 out of 3 years, 5.6 percent fail in all 3 years, and the remaining 4 percent pass in one out of 3 years. Among for-profit institutions, 41 percent pass in all 3 years, 14.6 percent pass in 2 out of 3 years, 27 percent fail in all 3 years, and the remaining 17 percent pass in one out of 3 years.¹²

Using data from the 2009 and 2010 cohorts, we can also examine within-institution correlations between 3-, 4-, and 5-year LRRs: 98 percent of public and nonprofit institutions and 90 percent of for-profit institutions would have the same status 3 and 5 years after repayment entry. When weighted by cohort balances at repayment entry, over 99 percent of public and nonprofit institutions and over 96 percent of for-profit institutions do not change status when outcomes are measured at 5 versus 3 years after repayment entry.

12. What about the students who attended failing programs in the years before any sanctions are implemented or outcomes are disclosed?

Answer: While not the focus of this paper, a complete accountability policy should also address how to assist students that enroll in programs that are later deemed ineligible for federal aid. For instance, automatically qualifying such students for federal student loan discharges and restoring their Title IV eligibility so that their time spent at the failing program does not count against their lifetime eligibility limit are two possible policies.

13. What about accounting for spending on instruction and other student supports or other accountability metrics that have been proposed?

Answer: Many of the accountability metrics that have been discussed are available as options in our online visualization tool (<https://higheredaccountability.com>).

¹¹ These data were provided by the Senate HELP committee to the authors. The data include each cohort’s balance at repayment entry, their balance 3 years after entering repayment, and the balance of borrowers whose loans are in forbearance, in-school or military deferment, or who have died or become disabled 3 years after entering repayment (“DDIS” borrowers). It is missing the initial balance DDIS borrowers. Thus, we cannot recreate the exact loan repayment rate from our main proposal. Instead, we subtract the balance of DDIS borrowers at year 3 from both the initial balance of the full cohort and the year 3 balance of the full cohort.

¹² In contrast to the set of schools that we focus on in our main analyses, we do not exclude for-profits that closed after 2016 from these analyses because repayment rates are measured before any borrowers would have received loan discharges through Borrower Defense to Repayment or Automatic Closed School Discharges.

References

- Barr, Andrew C., Kelli A. Bird, and Ben Castleman. 2019. "The Effect of Reduced Student Loan Borrowing on Academic Performance and Default: Evidence from a Loan Counseling Experiment." EdWorkingPaper: 19-89.
- Dunlop, Erin. 2013. "What do Stafford Loans Actually Buy You? The Effect of Stafford Loan Access on Community College Students." CALDER Working Paper 94.
- Marx, Benjamin M. & Lesley J. Turner. 2019 "Student Loan Nudges: Experimental Evidence on Borrowing and Educational Attainment." *American Economic Journal: Economic Policy* 11 (2): 108-141.
- Wiederspan, Mark. 2016. "Denying Loan Access: The Student-Level Consequences When Community Colleges Opt Out of the Stafford Loan Program." *Economics of Education Review* 51: 79–96.