Towards a framework for accountability for federal financial assistance programs in postsecondary education

Jordan D. Matsudaira
Columbia University

Lesley J. Turner
Vanderbilt University

The Brookings Economic Studies program analyzes current and emerging economic issues facing the United States and the world, focusing on ideas to achieve broad-based economic growth, a strong labor market, sound fiscal and monetary policy, and economic opportunity and social mobility. The research aims to increase understanding of how the economy works and what can be done to make it work better.

B | Economic Studies
at BROOKINGS

Contents

ABOUT THE AUTHORS

Jordan Matsudaira is an Associate Professor of Economics and Education Policy at Teachers College, Columbia University.

Lesley Turner is an Associate Professor of Economics at Vanderbilt University.

STATEMENT OF INDEPENDENCE

ACKNOWLEDGEMENTS

Introduction

The federal government invests significant taxpayer resources each year to support students' access to and success in postsecondary education. The main source of this support comes from federal student financial aid programs created under Title IV (TIV) of the Higher Education Act. In addition to these public funds, in many cases students and their families face significant out-of-pocket costs for postsecondary education. It is clear some of the dollars are not well spent, and, as a result, federal assistance programs designed to promote upward mobility and economic success instead do harm by subsidizing the exploitation of students' time and resources. Even very good investments may have substantial downside risks, and even the best higher education programs will have some students who fail to graduate, do not obtain employment, or struggle to repay their student loans. While eliminating all financial risk is not a realistic goal, it is irresponsible and imprudent to ignore the predictable harm of subsidizing students' enrollment at programs or institutions where adverse financial outcomes are expected or the norm.

In this paper, we outline a framework for an accountability system that would link federal financial aid eligibility to students' post-college outcomes and provide estimates of the benefits such a system might yield for students. We make the case for a conservative first step towards a comprehensive accountability system that sets minimum eligibility standards for programs offered by Institutions of Higher Education. These standards are based on two direct measures of student outcomes – earnings and loan repayment – with thresholds motivated by commonsense categorizations of unacceptable performance. In essence, the standards we propose would classify programs as eligible to participate in federal financial aid programs if either a) former students are able to make minimal, literally any, progress towards repaying what they borrow to attend; or b) a majority of former students find employment with earnings high enough to leave them no financially worse off than individuals with lower levels of education (high school graduates for those pursuing undergraduate credentials, and bachelor's degree holders for those pursuing graduate degrees) after accounting for the money they paid to attend college.

Background and motivation

Too many students are poorly served by their postsecondary education. The federal government regularly subsidizes college attendance at programs with a history of failing students, and institutions face few consequences if their students consistently struggle to repay their loans or find jobs that can support them and their families. For example, while

more than 1.23 million Direct Loan borrowers defaulted on their student loans in 2019, only 15 institutions were subject to accountability measures restricting Title IV eligibility to schools with high default rates.[1] As a result, many students waste time and money in educational programs where existing data suggest they would be better off having not enrolled (Cellini and Turner 2019; Gaulke et al. 2019) or enrolling elsewhere (Cellini and Chaudhary 2014; Armona et al. 2020; Cellini et al. 2020).

While existing policy does too little to protect students, past legislative and regulatory accountability efforts have been successful in shutting down or limiting the operation of low performing institutions and programs, and as a result, shifting students into better programs and improving post-college options. Sanctions based on cohort default rate (CDR) violations in the early 1990s led to large scale losses of federal student aid eligibility and closures for institutions where former students defaulted on their loans at high rates (Darolia 2013; Looney and Yannelis 2019; Cellini et al. 2020). Most students who would have enrolled in sanctioned schools were shifted into local community colleges, had lower student loan debt, and defaulted on their loans at lower rates (Cellini et al. 2020). The more recent Gainful Employment (GE) regulations, while never fully implemented, still may have had 'threat effects' that led some low performing programs to close (Kelchen and Liu 2019).

GE applied to non-degree programs and programs in for-profit institutions. While this made the rule a valuable tool for focusing on a sector and programs where poor outcomes were especially prevalent, poor performance was not limited to programs in institutions with for-profit tax status and/or non-degree programs. With the fall of many for-profit behemoths (e.g., ITT Tech and Corinthian Colleges), conversions of other for-profits into nonprofit institutions (e.g., Remington Colleges in 2010, Keiser University in 2011, Herzing University in 2015), and changes in how for-profits interact with students including through partnerships with public institutions (e.g., the Purdue and Kaplan partnership through Purdue Global, and the Arizona State and Ashford University partnerships), tax status may be a less useful proxy for sectors in which students face the most risk.[2] Moreover, the GE regulations were rescinded in 2019, with the Department of Education (ED) arguing that increased transparency around student earnings and borrowing outcomes through program level data in the College Scorecard made the rule unnecessary. However,

. . .

1.   See: https://www.ed.gov/news/press-releases/national-federal-student-loan-cohort-default-rate-continues-decline.

2.   For-profit and nonprofit entities in other markets also have been shown exhibit similar behavior and convert from nonprofit to for-profit status in response to financial incentives (e.g., Duggan 2000; Brickley and Van Horn 2002; Gaynor and Vogt 2002; Silverman and Skinner 2004; Capps, Carlton, and David 2020)

there is little evidence that outcome transparency alone will have substantial effects on where students, especially those from disadvantaged backgrounds, enroll in college (Blagg et al. 2017; Hurwitz and Smith 2018).

Accountability goals and metrics

An accountability system that governs federal student aid dollars for higher education could be used to advance many possible goals. For example, a system could be designed to promote more efficient allocation of public funds to higher value-added programs or to encourage institutions to improve the quality of their programs. These are important objectives but involve difficult measurement and design issues to avoid unintended consequences. Perhaps most challenging amongst these is how to appropriately measure the "value-added" of institutions or programs in a broad way that captures both the social and individual benefits, as well as the economic and non-economic benefits of the educations provided. We argue that a prudent first step would target a more modest goal: "do no harm." That is, we propose setting minimal standards that programs and institutions should meet in order to qualify for participation in federal student aid programs with the goal of protecting students from financial harm.

Proposed metrics

The standards we propose operationalize common-sense minimal expectations that students (1) should not borrow to attend a school or program where there is little reason to expect they can repay, and (2) should not attend a school or program if they will not have higher earnings than they would had they not attended, after accounting for what they paid. In short, we propose that IHE programs be eligible for federal financial aid programs if they satisfy *either* of two criteria:

1. The program has a federal loan repayment rate greater than zero, where the repayment rate equals the fraction of a cohorts' total loan balances measured at the date the cohort enters repayment that has been repaid 3 years later (the rate is negative when the cohort's balance grows); or

2. Most former students have net earnings exceeding a) median earnings of high-school graduates (in the case of undergraduate programs), and b) median earnings of bachelor's degree holders (in the case of graduate programs), where both comparisons are based on workers in the institution's state. A former student's net

earnings equal their annual earnings 3 years after exiting the program minus the out of pocket tuition and fees paid to attend the program (amortized to represent an annual payment).

The accountability system we outline would focus on program-level performance. This is in contrast to current policies which apply consequences to the institution as a whole based on institutional outcomes. We argue that a more targeted approach is warranted because, as we show in Section 7, most programs predicted to fail to satisfy both metrics are offered by institutions where most students attend programs that are predicted to pass. A system based only on institution level outcomes will both sanction students in programs with acceptable performance and fail to protect students in poor performing programs since aggregate indicators can mask both good and bad performance.

Principles for accountability metrics

While other metrics could potentially be incorporated in an accountability system, the two proposed metrics and corresponding passing thresholds satisfy several important principles.

1. <u>The metrics are based on student outcomes that are valued *per se*.</u> Any metric-based accountability system will create incentives for institutions to improve their measured performance. As such, it's important that the improvement in the metrics is associated with progress towards the goals of accountability. Our goals are to prevent students from attending programs that produce earnings too low to justify the cost, and with debt they cannot repay—the precise measures that would screen eligibility for student aid. Other potential metrics, including completion rates or input based measures of quality (e.g., instructional spending), are less clearly associated with the outcomes that represent whether students benefit from attending such programs.

2. <u>Accountability metrics should be difficult for schools to manipulate.</u> An abundance of evidence shows that accountability systems that rely on easily manipulated metrics will inevitably result in gaming (i.e., working to improve the performance metric rather than improving student achievement or outcomes) and cheating (i.e., falsifying reported outcomes) at the expense of meaningful improvement. For example, while linking TIV eligibility to student loan cohort default rates (CDRs) in the late 1980s and early 1990s was quite effective at reducing student debt and improving student loan repayment (Darolia 2013; Looney and Yannelis 2019; Cellini et al. 2020), institutions have developed methods of reducing CDRs without

actually improving loan repayment outcomes in recent years (GAO 2018).[3] Outcomes that are reported and/or verifiable by third parties are more difficult to manipulate than those that can only be verified by institutions themselves. In the case of the proposed metrics, student loan repayment comes directly from the ED while student earnings come from ED data linked to IRS earnings data.

3. <u>The outcomes used to determine eligibility should be measured quickly enough that most students are protected from enrolling in low-quality programs.</u> There is a trade-off between how accurately metrics capture students long-run economic well-being and how much of a lag is built in before performance is measured and any consequences can be applied. Metrics based on outcomes measured with more time elapsed since students leave college likely better reflect performance but entail more risk that additional cohorts of students will be exposed to poor programs. Both of the metrics we propose involve outcomes measured three years after students have left a program, a point in time when earnings and loan repayment outcomes have been shown to be broadly representative of such outcomes over the longer run (Chetty et al. 2017; Chou, Looney, and Watson 2017).[4]

4. <u>Metrics and thresholds are simple and easy to understand for institutions and prospective students.</u> Using metrics that are readily understood and viewed as

. . .

3. [3] GAO (2018) finds that some schools focus on improving their CDRs by encouraging students to put their loans in forbearance during the 3-year window that counts towards a school's performance. Examples of cheating in response to other accountability-based incentives include the Atlanta Public Schools cheating scandal (https://www.nytimes.com/2015/04/02/us/verdict-reached-in-atlanta-school-testing-trial.html) and numerous instances of higher education institutions reporting incorrect data to the U.S. News and World Report to improve their rankings (e.g., https://www.forbes.com/sites/christopherrim/2019/07/26/uc-berkeley-removed-from-us-news-college-rankings-for-misreporting-statistics/#3a734d437578).

4. [4] Among a cohort of students enrolled in two-year colleges and "non-elite" four-year colleges, Chetty et al. (2017) find that students' rank in the earnings distribution stays is relatively constant between the ages of 25 and 36. Students at more elite colleges experience steep increases in earnings ranks between 25 and 30, and then stabilize. The data are not perfectly comparable to those that would be used in the proposed accountability metrics, but provide some evidence that for non-elite institutions measuring earnings 3 years after program exit, when most students will be near or over the age of 25, will provide an accurate ranking of students' labor market outcomes across programs over the longer-run. The Chetty et al. (2017) data are based on older cohorts of students enrolled between 1999 and 2000 while near the age of 20. By construction this omits older, independent students, who comprise a larger share of "non-elite" colleges. It is not clear whether these patterns are representative of all students at such institutions. With respect to loan repayment, institutional 3-year loan repayment rates are highly correlated with long-run repayment outcomes (Chou, Looney, and Watson 2017). Using supplemental data from the 2009 and 2010 repayment cohorts (provided by the Senate HELP Committee to the authors, upon request), we estimate that over 95 percent of institutions would have the same repayment rate status (e.g., pass or fail) at 5 years after repayment entry as they would at 3 years and, when weighted by cohort balances at repayment entry, more than 99 percent of institutions would have the same status at 3 and at 5 years post-repayment entry. We provide additional details of these analyses in Appendix C.

measures of quality by students and institutions increases the legitimacy of the accountability system. It also opens the possibility that producing and disseminating information in the metrics can have the added benefit of better informing students' program choice. Finally, as the thresholds for passing each metric represent a clear lower bound for the minimal expected outcomes to ensure a student is not harmed, they avoid being perceived as arbitrary.

5. <u>The proposed metrics provide separate measures of financial well-being and compensate for weaknesses in the other.</u> Both metrics are valuable proxies for former students' financial well-being, and each individually could form the basis of a reasonable accountability system. There are benefits, however, in combining the two metrics by linking eligibility to participate in TIV programs conditional on passing *either* metric. For example, administrative earnings reported to the IRS may not reflect the full earnings of individuals in some occupations, such as cosmetology, due to underreporting of tipped income. Additionally, some programs may generate low earnings for students, but students may fully anticipate this and have resources in place—such as parental support—that will allow them to repay their loans. In both cases, loan repayment rates serve as an alternative measure of students' financial well-being. Similarly, repayment rates alone might not be a perfect measure of economic well-being if they primarily reflect administrative hurdles to repayment for some students.

Measurement details for metrics

In this section, we describe the construction of the two proposed accountability metrics and the thresholds delineating satisfactory performance. We note where additional data collection would be required for the ideal construction of these metrics.

Net earnings premium (NEP) metric

The first metric evaluates whether students are financially better off having attended a program of study than they would have been if they hadn't enrolled. This is operationalized by grouping students into "exiter cohorts" consisting of all students leaving a program—including completers and non-completers—in a particular year (excluding those currently enrolled in a different institution or program of study, and those who are disabled or deceased). For each cohort, the net earnings premium equals median cohort earnings less

out-of-pocket (OOP) costs of attending the program (amortized to represent an annual payment over a given period) relative to their "counterfactual earnings" measured by the median earnings of individuals with lower education levels (i.e., high school graduates for undergraduate programs or bachelor's degree recipients in similar fields for graduate programs) in the same state:

$$NEP = (median\ cohort\ earnings) - (median\ cohort\ OOP\ costs)$$
$$- (median\ counterfactual\ earnings)$$

Median cohort earnings are measured three years after program exit for all former students in the exiter cohort with positive earnings over the course of the year.[5] In contrast to the program-level earnings information in the College Scorecard data, cohort earnings in the NEP metric include the earnings of non-completers. This is critical, in our view, to ensure the metric creates incentives for programs to increase graduation rates and not let programs 'off the hook' for failing to do so.[6]

From this measure of students' earnings, we subtract a measure of the annual financial and the annual opportunity cost of attending the program, respectively. Our proposed measure of students' financial cost of attending a program is the cumulative amount that students pay towards tuition and fees "out of pocket" over the course of their enrollment in a program, excluding grant and scholarship aid, but including loan dollars applied to tuition and fees.[7] Although living expenses are included in cost of attendance calculations by institu-

. . .

5. While a case could be made for including students with zero earnings in the median earnings calculation, doing so would effectively penalize programs with high rates of non-employment, where the non-employed include those who want to work but are unable to (i.e., unemployed workers) but also those who are not working by choice (i.e., those out of the labor force). While it may be appropriate to treat all nonemployment as a negative outcome in an accountability system limited to programs whose goals are tightly connected with securing employment for students, since most long-term nonemployment is comprised of individuals not in the labor force, it seems more appropriate (albeit conservative) to use the median earnings of workers in a comprehensive accountability system.

6. At the same time, cohorts could exclude student whose attachment to the institution is de minimis to screen out students that may not have intended to take more than a few credits.

7. An argument could be made for including all costs incurred by society of a student's education, including those paid by federal, state, and local governments, and an accountability policy with broader aims to increase the efficiency of educational expenditures might better be based on such a metric. Here, we maintain our focus on protecting students from financial harm, and only consider costs paid by the student. Even programs where the median student faces no out-of-pocket costs will still need to provide enough value that most former students earn more than the reference group. Additional benefits of our approach are the incentives provided to institutions to 1) reduce net tuition by offering greater institutional grant aid and 2) not reduce low-income students' access, as these students would face lower out-of-pocket costs due to support from federal and state need-based grants. Ideally, we would collect information on tuition

tions and ED, we exclude these from our measure both to make our calculations conservative and avoid the conceptual difficulties of attempting to measure the 'incremental' living costs involved in postsecondary enrollment. We convert the measure of cumulative out of pocket spending on tuition and fees to a yearly cost measure by amortizing the costs over 20 years for undergraduate programs and 25 years for graduate programs. These amortization periods are consistent with the repayment horizon in federal income-based repayment plans and align with the long working careers over which most students can expect to receive higher earnings based on skills they acquire through higher education.[8]

Our measure of opportunity cost captures what students would have earned had they not pursued enrolled in higher education (i.e., their counterfactual earnings). For undergraduate programs, this measure is based on the median earnings of high-school graduates or GED holders aged 25 to 34 in the state where the institution is located. Median earnings for former graduate students are benchmarked against the median earnings of recipients of bachelor's degrees in the same broad field, aged 25 to 34, and in the state where the institution is located.[9] In practice, this is estimated using the American Community Survey based on all individuals with the relevant credential (and field, for graduate students), including those who report zero (or negative) earned income over the past year or who report being unemployed when they were surveyed.[10] By including both unemployed individuals and those not in the labor force, this measure builds in a natural responsiveness of our

. . .

and fees and grants and scholarships separately to allow the out-of-pocket cost component of the NEP metric to be negative to reward institutions that invest in affordability. Doing so, however, would likely require modifications to reporting on the 1098-T form, or new data reporting requirements.

8. [8] The recently rescinded Gainful Employment regulations used longer amortization periods for programs that required more years to finish. One might justify such a structure by the fact that earnings growth tends to be steeper for longer-term programs, and so earnings measured at a fixed number of years post-separation might understate the lifetime earnings for students in longer relative to shorter programs. Amortizing costs over more years might partially offset such a bias, but it is a coarse adjustment that may not be well targeted in all cases (e.g., due to heterogeneity in wage growth across programs) and in practice does not significantly alter the set of programs deemed to pass the NEP. We therefore prefer the simpler structure of a common amortization period for all undergraduate programs and for all graduate programs.

9. We measure median bachelor's degree earnings in five broad fields: Arts and Humanities; Education and Public Service; Agriculture, Consumer Services, and Trades; Business and Social Sciences; and STEM and Health fields. Median counterfactual earnings range from about $37,000 for Arts and Humanities bachelor's degree graduates to nearly $53,000 for STEM and Allied Health bachelor's degree graduates. Appendix B provides additional detail on these broad field of study categories and how we match them to field of study categories at the 2-digit classification of instructional program code level for graduate programs.

10. Although the majority of individuals not employed are not actively looking for work, again in the interests of being conservative, we base the counterfactual earnings measures on all individuals rather than only those in the labor force.

metric to the business cycle. Using state-specific counterfactual earnings accounts for differences in economic conditions across states.[11]

Loan repayment rate (LRR) metric

Our second accountability metric directly captures how well students who borrow to attend a program are able to repay their debt. For a given "repayment cohort" comprised of all students entering repayment on their federal loans in a particular year, the repayment rate is equals one minus the ratio of the cohort's outstanding balance measured at the end of the third year after entering repayment relative to their total balance at repayment:

$$LRR = 1 - \frac{(Balance\ in\ year\ 3)}{(Balance\ at\ repayment)}$$

The balances of borrowers who are in school, have a military deferment, or not required to repay their loans for other reasons (e.g., death or disability) as of their third year after entering repayment are excluded from both the initial cohort balance and the balance in year three.[12] The passing threshold we set for this measure is that it should be positive, which is equivalent requiring that by three years after entering repayment, the cohort as a whole has reduced their loan balance by at least $1. This operationalizes the conservative standard that federal dollars should not subsidize students attending programs where past students, as a collective, failed to make *any* progress in paying off their loans.

There are possible alternatives to this dollar-based repayment rate metric. For example, the College Scorecard reports a "person-based" repayment rate that is similar in spirit to the "dollar-based" measure we propose. The person-based metric calculates the fraction of borrowers in a cohort that have paid down at least $1 (and are not in default on their loans), and a parallel construction to our dollar-based measure would be to consider a program passing if the median student has made progress paying down their loan—that is, the repayment rate is greater than 50 percent. In our view both measures are sensible, but the dollar based metric we propose provides stronger incentives for institutions to be accountable for all taxpayer dollars that their students borrow, focus on encouraging repayment

. . .

11.   For online institutions, we use national median earnings for the same reference population.

12.   Specifically, we first identify borrowers who have an in-school or military deferment or have died or become disabled as of their third year after entering repayment. For this group, we calculate their initial balance at repayment entry and their year three balance. We subtract their initial balance from the denominator of the LRR metric and their year three balance from the numerator.

among students with higher balances, and take steps to reduce the amount students need to borrow.[13]

Programs that do not have a repayment rate because their institution does not participate in federal loan programs will be evaluated based on the net earnings premium metric alone. The alternative of classifying these programs as automatically eligible would create undesirable incentives for schools to opt-out of participating in federal loan programs or stop packaging student loans, decisions which have been shown to be detrimental to students, on average (Dunlop 2013; Weiderspan 2016; Marx and Turner 2019). Furthermore, students who invest their own money and time and are supported by other sources of federal student aid but cannot access federal loans due to institutional decisions should still be protected.[14]

Measurement issues and new reporting requirements

Ideally, an accountability system would apply to programs in the way that students interact with them—at the level of specific degree offerings, perhaps split by whether the program is predominantly online or in-person. In practice, data limitations force some compromises. Some are driven by current data limitations that could be addressed, while others driven by practical considerations around the minimum number of students that need to be enrolled in a program to produce reliable metrics and protect students' rights to privacy. We propose that the net earnings premium and loan repayment rate metrics be produced at the institution (defined by the Integrated Postsecondary Education Data System or IPEDS 'unit identification number') by credential level by field of study level, where program of study is based on 2-digit Classification of Instructional Programs (CIP) codes.[15]

. . .

13. While we do not have person-based repayment rates available at the program level, at the institution level the two metrics are highly correlated (r=0.793). Nonetheless, the person-based metric has greater bite: while 71 percent of schools have a positive (passing) dollar-based metric, only 42 percent of schools pass a person-based threshold that would require the median borrower to have reduced their balance by $1 (both measured three years after repayment entry). Appendix Figure A1 shows the relationship between the two measures.

14. To evaluate programs in schools that do participate in federal student loan programs but have a small number of borrowers, several cohorts of borrowers can be "rolled-up" into a single cohort. This would be similar to the "average rate formula" for calculating institutional cohort default rates when a cohort has fewer than 30 borrowers entering repayment in a single year.

15. Categorizing institutions based on the IPEDS unit ID rather than the OPEID will require institutions to report enrollment based on unit ID numbers in administrative data systems like the National Student Loan

Institutions should also be required to delineate between programs that are predominantly offered online versus in-person, given well documented differences in outcomes for students in partially or fully online programs (Bettinger et al. 2017; Hart et al. 2018; Hoxby 2019).

Defining programs at the 2-digit CIP code level is coarser than desirable but allows a much higher fraction of programs to be covered by our accountability scheme. We estimate that 53 percent of 2-digit CIP programs (approximately 80 percent of 4-digit CIP programs), containing an estimated 9 percent of exiting students, would be missed if outcomes were measured at the 4- versus 2-digit CIP level.[16] Furthermore, programs' NEP performance at the 4-digit CIP level almost never differs from performance when programs are aggregated to the 2-digit level.[17] Nonetheless, for the purpose of providing useful information to students and their families, we suggest that our metrics also be reported at the 4-digit level for disclosure purposes where possible.

Other data reporting modifications may facilitate calculation the NEP metric. The IRS currently collects amounts paid for qualified tuition and related expenses at eligible educational institutions on an annual basis for all postsecondary students. This information could be reported on an aggregated basis via data sharing agreements between the Treasury Department and ED, just as earnings data are made available for the College Scorecard.[18] Alternatively, Congress could require that information be shared with ED, or via a data sharing agreement between the Treasury Department and ED. Requiring institutions to report the number of students in each program who are predominantly online, or to treat online programs as separate programs, would provide students with better information to inform school and program choice. For programs in which a majority of students are

. . .
Data System (NSLDS), but is important to align the metrics with the way in which institutions present themselves to students (as opposed to how their administrative functions are organized).

16. This estimate is based on a comparison between the number of programs that have a valid repayment rate (reported at the 2-digit CIP level) but are missing earnings data (reported at the 4-digit CIP level). For a listing of 2- and 4-digit CIP codes, see https://nces.ed.gov/ipeds/cipcode/browse.aspx?y=55.

17. Specifically, for programs defined at the 2-digit CIP level that have more than one component program at the 4-digit CIP level with non-missing earnings data, 93 percent have no variation in NEP performance at the 4-digit CIP level (i.e., all component 4-digit CIP programs pass if the 2-digit CIP program passes or all component 4-digit CIP programs fail if the 2-digit CIP program fails). We cannot evaluate differences in performance on the LRR metric between the 2- and 4-digit CIP code level because the data we have available to construct the LRR metric is reported only at the 2-digit CIP level.

18. Institutions file 1098-T information reports to the IRS that report qualified education expenditures for each student that include tuition, fees, and course materials (See: https://www.irs.gov/instructions/i1098et).

online, the NEP metric should use the national median rather than earnings that are specific to the institution's location.

Simulating program performance: Data and methodology

In this section, we describe the data and methods used to simulate the effect of using the proposed accountability metrics to determine TIV eligibility. Programs are defined at the field of study (2-digit CIP) by credential level by institution (using six-digit OPEID) level. Program-level loan repayment rates were calculated directly by ED using NSLDS data and are based on the loan balances of individuals who entered repayment in FY 2016.[19] While ED could also calculate the net earnings premium (NEP) metric directly (subject to the changes proposed in Section 4.3), we must make assumptions to approximate program performance based on existing data. We provide an outline of this process here and additional details in Appendix B.

To measure program-level median earnings, we rely on College Scorecard data, which includes the median earnings of students who completed a program (measured at the 4-digit CIP code level) in the 2014-15 and 2015-16 award years. To convert the earnings in the College Scorecard into the earnings measure used in the NEP metric, we need to account for three differences. First, that Scorecard earnings are reported as of the first calendar year after graduation, while we propose measuring earnings three years after exit. We use estimates of annual earnings growth rates by major and credential level from the American Community Survey (ACS) to make this adjustment. Second, our earnings measure would include both completers and drop-outs, while Scorecard earnings are limited to program completers. We use average earnings by completion status for borrowers to estimate the 6-year completion rate of TIV borrowers and the ratio of completer earnings to non-completer earnings at the institution-level. We apply these institution-level ratios to year-three earnings calculated in our first step for all programs within the same institution. In essence, observed earnings are adjusted downward more when the earnings premium for completers relative to non-completers is larger, and when the rate of non-completion is higher. Finally, because Scorecard earnings are reported at the 4-digit CIP code level, while our proposed metrics are calculated at the 2-digit level, in the third step, we calculate the

. . .

19. These data were provided to the Senate HELP Committee by ED and shared with the authors upon request.

weighted average of the measure for all 4-digit programs within a given 2-digit CIP code, with weights equal to the estimated number of annual TIV exiters for each 4-digit CIP program.

To estimate program-level out-of-pocket expenditures for undergraduate students, we calculate institution-level net tuition paid per student and multiply this amount by the average number of years undergraduate students remain enrolled in the institution (based off of persistence and completion rates from institution-level College Scorecard data). We take the ratio of this measure to median undergraduate student debt at the institution-level (also from College Scorecard data), which provides an approximation of the share of out-of-pocket costs covered by median debt. We then apply this ratio to program-level median debt for undergraduate programs to generate an estimate of program-level out-of-pocket costs for undergraduate programs. We approximate graduate program out-of-pocket costs with median program-level debt.[20]

Finally, we estimate is the number of TIV exiters for each program using information on the number of program completers (captured in IPEDS and included in the College Scorecard program data). In total, we estimate that approximately 7.6 million TIV recipients left college in 2016. Approximately 9 percent of all (estimated) TIV exiters attend programs that have neither a repayment rate nor a measure of net earnings; characteristics of these programs can be found in Appendix Table A1.

Simulated performance on the proposed metrics

How does performance on these two metrics vary across the programs likely to be covered by an accountability scheme based on net earnings and loan repayment? In this section, we illustrate and analyze the correlates of performance on each metric, and then discuss the number and characteristics of programs that are likely not to meet the passing thresholds for both metrics.

We summarize program performance in two ways. First, we present 'raw' descriptive statistics on the percentage of programs that we predict to fall below the thresholds for each metric and the combination of both metrics. Second, to get a sense for how many students are in such programs, we also present estimates weighted to represent the number of TIV

. . .

20.   This is because we do not have comparable information on institutional persistence and completion rates or median debt at the graduate-level.

recipients that would 'exit' annually from all programs (out of an estimated 7.6 million students in total). We present broad descriptive statistics for each metric first, and then illustrate performance across both measures by credential level, sector, and major category.

Net earnings premium metric

The net earnings premium (NEP) measures the difference between median program earnings (net of out-of-pocket costs) and the earnings of a reference group meant to proxy their "counterfactual earnings" or what they would have earned if they had not enrolled in the program. For undergraduate programs, the reference group is high school graduates from the state where the institution is located and for graduate programs, the reference group is recipients of bachelor's degrees in similar fields.

Program NEPs vary dramatically. Across the 22,669 programs for which we have sufficient data to calculate this metric, the average NEP (i.e., the average of the difference between the median student's net earnings and median reference group earnings across programs) is $12,817 ($11,764 when we weight the data to be representative of all exiters). We estimate that, in a given year, 15 percent of all exiters attended programs with negative net earnings premia, with a larger share in graduate programs (24 percent) than in undergraduate programs (13 percent). Despite the larger share of graduate programs with negative net earnings, average net earnings premia rise with credential levels. Weighted by the number of annual exiters, average net earnings range from $6,917 for certificate programs, $7,827 for associate's programs, $17,715 for bachelor's programs, to $23,855 for doctoral programs.[21]

Table 1 depicts the fraction of enrollment and the fraction of programs with negative net earnings premia by credential level and institutional control. Panel A shows the estimated percentage of all annual exiters enrolled in these programs to provide the best sense for the number of students that might be affected by a policy restricting program eligibility for aid based on the NEP metric. Approximately 15 percent of all students attended a program with a negative NEP. Among undergraduate programs, certificate and associate degree programs have by far the most students enrolled in programs with negative earnings premia at 23 and 17 percent, respectively. Only about 3 percent of students from bachelor's degree programs attended a program with a negative NEP. Among graduate programs, where we compare net earnings to a higher reference group level of earnings, about one quarter of

. . .

21.   For completeness, the average NEP is $14,384 for graduate certificate programs, $10,788 for master's degree programs, and $18,792 for first professional degree programs.

enrollment in certificate, master's, and first-professional degree programs are in programs with a negative NEP. Among doctoral degree programs, only about 7 percent of students attend such programs.

Table 1 also illustrates several stark differences across sectors. About 30 percent of students leaving for-profit institutions came from programs with negative NEPs, compared to 12 and 14 percent in public and private, nonprofit programs respectively. The difference is driven by certificate programs and graduate programs (excluding doctoral programs), with 36 percent of certificate-seeking for-profit students coming from failing programs compared to 26 percent of certificate-seeking students from private nonprofits, and 16 percent of those from public programs. Among graduate certificate, master's, and first professional degree programs, differences across sectors are even more pronounced: in each sector about two-thirds of for-profit students attended 'failing' programs, whereas in public and private nonprofit programs the share ranges from 11 to 33 percent of students. Because the for-profit sector is a small share of overall graduate program enrollment, most students attending graduate programs with negative NEPs attended nonprofit and public institutions.[22]

We provide additional information on the variation in NEP performance within each credential level by field in Panel A of Figures 1A to 1G (Panel B contains information on program performance on the LRR metric, which we discuss in the following section). The top five fields of study (in terms of TIV exiters) are shown for each credential level by institutional control (enrollment shares are shown in parentheses).[23] The range of each box represents the 25th through 75th percentiles of net earnings for programs in the sector, field, and credential level (weighted by the estimated number of exiters in each program). The line that falls within the box represents the 50th percentile or the level of net earnings at which half of all (exiter-weighted) programs have worse outcomes and half of all programs have better outcomes The "whiskers" extend to the 5th percentile to the left and the 95th percentile to the right. The figures contain a wealth of detail on performance for the five most popular fields at each credential level, but some broad generalizations are possible.

. . .

22. We estimate that 72 percent of all exiters from failing undergraduate and graduate programs attended programs in public or nonprofit institutions. For-profits account for the remaining 28 percent, despite containing less than 14 percent of all exiters.

23. We combine a handful of 2-digit CIP codes for the purpose of displaying performance for a greater share of programs. Services includes personal and culinary services and family/consumer sciences (CIP codes 12 and 19). CIS includes computer and information sciences and math and statistics (CIP codes 11 and 27). Social sciences includes psychology programs (CIP codes 42 and 45). For the performance of all programs within a given 2-digit CIP code, see Appendix Figures A3 and A4. For enrollment shares by 2-digit CIP for each credential level, see Appendix Table A2.

First, performance on the NEP metric varies dramatically within most fields of study and within sectors. At the sub-baccalaureate level, nearly every field of study contains passing and failing programs (judged by whether their NEP is positive or negative), as well as a wide distribution of net earnings premia within each sector. Among undergraduate certificate programs, programs in service-related fields (e.g., cosmetology, culinary services, and family and consumer sciences) fare the worst by far among popular fields. It is the only popular field and credential level where a majority of students attend failing programs. The strong sectoral differences in performance shown in Table 1 are driven in part by allied health programs, where the distribution of NEP performance is much higher for students in public programs than for those in private or for-profit programs.

Popular associate degree programs (Figure 1B) with negative NEP in the public and private sector are heavily comprised of liberal arts/ general studies programs, echoing previous research that finds much lower economic returns for students in such programs at community colleges (e.g., Liu et al. 2015; Bettinger and Soliz 2016; Turner 2016). The performance of these programs is important given that an estimated 46 percent of all TIV students who leave associate degree programs pursued a degree in this field. By in large, bachelor's degree programs with high enrollment have overwhelmingly positive NEP (Figure 1C). Of the top five fields (in terms of enrollment) only Biology programs in nonprofit institutions have more than 5 percent of enrollment in programs with negative net earnings.[24]

Figures 1D through 1G show again wide variation in graduate program net earnings premia both within and across credential level and program type. Across graduate certificate, master's degree, and first professional degree programs, those in the for-profit sector have particularly poor performance, and differences across sectors are particularly pronounced in the allied health field. Using bachelor's degree recipients are the reference group for estimating counterfactual graduate earnings raises the threshold for a positive NEP considerably: If we calculated graduate programs' NEP using the earnings of high-school graduates as their counterfactual earnings, nearly all graduate programs would have positive net earnings premia (see Appendix Tables A3 and A5).

Appendix Table A6 provides information on the characteristics of the 19,183 programs with positive (passing) and 3,486 programs with negative (failing) earnings premia, shown separately for undergraduate and graduate programs. Over 40 percent of undergraduate programs that have negative net earnings premia are in for-profit institutions while only 8

. . .

24. Appendix Tables A3 and A4 show average NEP and the percent of undergraduate programs with negative NEP using different reference groups for estimating counterfactual earnings and for different amortization periods for out-of-pocket costs.

percent of passing programs are in for-profits. Over 90 percent of failing for-profit undergraduate programs are undergraduate certificate programs. Public associate degree programs are also more prevalent among programs failing the net earnings premium metric than among passing programs (24 percent versus 15 percent, respectively). On average, after accounting for out-of-pocket expenses, students in failing undergraduate programs earn $3,700 less than a typical high school graduate. Among passing programs, the average NEP exceeds $15,000. The difference in net earnings premia between failing and passing undergraduate programs is driven by differences in earnings (approximately $36,500 per year for passing programs versus $17,400 for failing programs) and not by the other components of the metric. Failing programs actually have lower estimated total out-of-pocket costs than passing programs (approximately $21,100 versus $27,900) and are located in states with similar median earnings for high school graduates.

Still focusing on undergraduate programs, we also find substantial differences in the representation of fields among failing and passing programs. Agriculture, trades, and service programs (driven especially by personal and culinary service programs) make up 46 percent of all programs with negative net earnings premia compared to 17 percent of undergraduate programs with a positive NEP. Likewise, arts and humanities programs (including liberal arts and general studies) make up a greater share of failing programs (30 percent versus 18 percent of passing programs). Conversely, STEM and allied health programs are more prevalent in the group of passing programs than among failing programs (33 versus 11 percent of programs, respectively), as are business programs (15 percent of passing versus 5 percent of failing), and education and public service programs (18 percent versus 8 percent).[25]

At the graduate level, we find somewhat similar patterns in program performance, with for-profit programs making up a larger share of failing versus passing programs (7 versus 3 percent, respectively). However, overall differences both by sector and by field are more. Most graduate programs with negative net earnings offer master's degrees and are in public or nonprofit institutions but this is due to the relative prevalence of such programs rather than a higher relative rate of failure. Estimated out-of-pocket costs are substantially higher

. . .

25.   The third column of Appendix Table A6 displays characteristics of the bottom 1 percent of all undergraduate programs with respect to net earnings. For-profit institutions make up 66 percent of these schools, 23 percent are nonprofits, and the remaining 11 percent are public. Certificate programs are the predominant credential offered in the bottom 1 percent of programs, making up 74 percent of all such programs, with the vast majority in for-profit institutions. Average estimated out-of-pocket costs are similar in the bottom 1 percent of programs to all programs with negative earnings premia but median earnings are substantially lower similar (approximately $13,000 on average for bottom 1 percent programs versus $17,000 for all programs with negative earnings premia). Personal and culinary service programs make up 68 percent of undergraduate programs with net earnings in the bottom 1 percent.

than undergraduate costs and are almost twice as large for graduate programs in the bottom 1 percent of net earnings than for other graduate programs (approximately $102,000 versus $54,000 for passing and $57,000 for failing programs).

Loan repayment rate metric

Turning to the loan repayment rate metric, we have data for about twice as many 2-digit CIP programs (46,375) because the repayment rate data are reported at the 2-digit CIP level, so privacy suppression due to small numbers of students in a program is less common. Weighting by student exiters, the average repayment rate across all programs is 5.8 percent – this means that three years after entering repayment, an average cohort as a whole still had 94.2 percent of their original balance left to repay. It is worth pausing to emphasize how low this rate is: on a standard 10-year repayment plan, and ignoring interest, we would expect students to have paid about 30 percent of their loans by this point in time. Overall, 36 percent of students attended programs with negative repayment rates.

Repayment rates do not generally improve at higher credential levels, although similar to patterns of performance on the NEP metric, bachelor's degree programs have the best loan repayment outcomes. The average certificate-seeking student attended a program with a repayment rate of 4.9 percent (95.1 percent of original balance remaining). Average repayment rates are 2.3 percent (97.7 percent of balance remaining) for associate degree programs, 11.4 percent (88.6 percent of balance remaining) for bachelor's degree programs, 7.0 percent (93 percent of balance remaining) for graduate certificate programs, 5.3 percent (94.7 percent of balance remaining) for master's degree programs, 7.1 percent (92.9 percent of balance remaining) for doctoral programs, and 8.5 percent (91.5 percent of balance remaining) for first professional programs.

Panel A of Table 2 shows differences in the fraction of students in programs with negative repayment rates by credential level and sector. These range from a high of 51 percent for associate degree programs, to a low of 19 percent for bachelor's degree programs. The majority of for-profit students (56 percent) attend programs with negative repayment rates, compared to 34 percent of students from public programs and 29 percent of students from private, non-profit programs. In fact, outside of undergraduate certificate programs, the

majority of for-profit students attend failing programs.[26] The share of programs with a negative LRR is similar, albeit slightly lower, in the unweighted estimates (Panel B).

Panel B of Figures 1A to 1G show the distribution of loan repayment rates for the most popular majors within each credential level by sector. Again, we observe substantial differences in performance across programs within a given field by credential level by institutional control category, but also substantial overlap in performance across sectors. Across sectors, majors with the highest net earnings premia tend to have better LRR performance, but the correlation is imperfect suggesting that the NEP and LRR metrics indeed pick up different aspects of economic well-being. Poor performance on the LRR metric is much more common across sectors at the sub-baccalaureate level, even in fields where many programs show positive net earnings. Differences across sectors are particularly pronounced at the baccalaureate level and above, where for-profit programs' repayment rates are substantially worse across many of the high-enrollment fields.

The first and second columns of Appendix Table A8 show characteristics of the 13,341 undergraduate programs with positive (passing) loan repayment rates and the 3,601 undergraduate programs with negative (failing) repayment rates, while the fourth and fifth columns provide similar results for the 3,459 passing and 1,431 failing graduate programs. Only 5 percent of passing undergraduate programs and 1 percent of passing graduate programs are in for-profit institutions, while 12 and 6 percent of undergraduate and graduate programs with negative repayment rates are in this sector. Public institutions are also slightly over-represented in the set of failing undergraduate programs (67 versus 57 percent of failing and passing programs, respectively). Public associate degree programs are the most over-represented in the set of failing programs (36 percent of failing undergraduate programs versus 18 percent of passing programs), while public and nonprofit bachelor's degree programs are underrepresented, with public (nonprofit) bachelor's degree programs making up 18 (17) percent of programs with negative repayment rates and 31 (34) percent of undergraduate programs with positive repayment rates. At the graduate level, certificate programs in nonprofit and for-profit institutions are slightly over-represented while other credentials and sectors are relatively equally represented among failing and passing programs.

On average, undergraduate cohorts in schools with negative repayment rates saw their loan balances rise by 4 percent – from $1.09 to $1.12 million – in the 3 years after entering

. . .

26. Differences in loan repayment by sector and credential level are more skewed when reweighting by the initial loan balance of the repayment cohort (Appendix Table A7), suggesting that cohorts that accumulate larger per-borrower balances find it harder to keep up with loan repayment.

repayment. Cohorts from failing graduate programs saw even larger increases in their balances ($2.66 to $2.77 million, or 5 percent). In contrast, undergraduate programs with passing loan repayment rates paid down an average of 13 percent of their balance, going from $1.01 million to $0.89 million in outstanding student debt, putting these cohorts on track to repay their student loans in approximately 15 years. Passing graduate program cohorts saw their balances fall by 12 percent, from $3.03 to $2.65 million, on average.

We find some differences in the representation of fields among programs with positive and negative loan repayment rates compared to the representation of fields when performance is based on the NEP metric. STEM and allied health fields are underrepresented in undergraduate and graduate programs with negative loan repayment rates and graduate business programs are also slightly underrepresented. Security/protective services programs (included in the education and public services category) are overrepresented in both undergraduate and graduate programs with negative loan repayment rates. Other programs overrepresented in the set of programs failing the loan repayment rate metric include liberal arts/general studies undergraduate programs (included in the arts and humanities category) and graduate psychology and public administration programs (included in the business/social sciences and education and public services categories, respectively) are also overrepresented in programs with negative loan repayment rates.[27]

Which programs perform poorly on both metrics?

We propose that institutions should be required to pass either the net earnings premium metric or the loan repayment metric as a condition for TIV aid eligibility. In this section, we discuss the characteristics of programs that have both a negative earnings premium and a negative loan repayment rate, and the share of TIV exiters across sectors and credentials in such programs.

Figures 2A to 2G show scatter plots of the estimated NEP and LRR metrics for each of the programs in our data, separately by credential level and institutional control (panels A, B, and C within each figure). Each marker represents a single program's performance and

. . .

27. The third and sixth columns in Appendix Table A8 display the characteristics of the bottom 1 percent of all undergraduate and graduate programs with respect to LRR performance. Cohorts in undergraduate (graduate) programs that fall in the bottom 1 percent of loan repayment rates saw their balances increase by 11 (14) percent in the 3 years after repayment entry. Among undergraduate programs, business and public administration are overrepresented in the bottom 1 percent while allied health programs are underrepresented. For graduate programs, security/protective services are overrepresented while allied health and education programs are underrepresented in the bottom 1 percent of loan repayment rates.

markers for programs in the top five majors (in terms of enrollment) are depicted in colored shapes described in the legend. The pass-fail thresholds for each metric, that is a positive net earnings premia and a positive loan repayment rate, are indicated by the lines at zero on each axis so that programs failing both metrics will be in the lower left quadrant of each figure.

Table 3 provides a high-level summary of performance on both metrics, showing the fraction of exiters and programs that either fail both metrics, pass one but not the other, or pass both metrics.[28] As shown in Panel A of Table 4, we estimate that 9 percent of all TIV recipients attend programs that fail both metrics. Approximately 20 percent of student in for-profit institutions attend such programs compared to 7 and 8 percent of students in public and nonprofit institutions, respectively. Enrollment in programs failing both metrics is by far most prevalent in graduate programs (excluding doctoral programs), with between 18 and 20 percent of students attending programs with negative repayment rates and earnings premia. In contrast, only 1 percent of bachelor's degree seeking students and 3 percent of doctoral degree students attend programs that fail both metrics. Sub-baccalaureate programs have failure rates that fall between these extremes, with 13 and 11 percent of undergraduate certificate and associate degree seeking students attending failing programs, respectively.

The second and third columns of Table 3 provide insight into how the bite of a system based on our proposed metrics would differ if it were based only on one of the two metrics instead, while the fourth column provides insight into the share of programs that would retain TIV eligibility if programs were required to pass both metrics instead of one or the other. More students attend programs with negative repayment rates (23 percent) than negative net earnings premia (18 percent), meaning that more programs would be affected by a system in which programs needed a positive repayment rate to participate in TIV aid programs. For certificate programs, however, there are a substantial number of programs that pass based on positive repayment outcomes despite negative net earnings (23 percent of undergraduate certificate programs and 13 percent of graduate certificate programs). This is also somewhat true of associate degree programs and master's degree programs, with 7 and 12 percent of programs passing based on loan repayment rates alone. At the bachelor's and doctoral degree levels, the net earnings premium metric is rarely binding, and very few programs fail it. On the other hand, in a system based on the repayment metric alone, 13

. . .

28.   There are a small number of programs in schools that do not participate in federal loan programs (N = 445). For the purpose of the analysis presented in Table 3, we classify such a program as having a positive LRR if it has a positive NEP and having a negative LRR if it has a negative NEP.

percent of BA programs would fail, and between 17 percent of all doctoral degree programs would fail.

In Panel B of Table 4, we provide estimates of the number of students that might be affected by an accountability system based on these two metrics and these students' monetary costs of enrolling in programs with negative net earnings and repayment rates. We estimate approximately that in the absence of any accountability policy, 670,000 students per cohort enroll in a program with both a negative NEP and a negative LRR. These students pay total of $17.1 billion in estimated out-of-pocket costs and leave their programs with $19.4 billion in aggregate debt. Three years after exit, their outstanding debt is $800 million higher than when they entered repayment. Denying eligibility for federal aid would likely lead students in affected programs to transfer to better programs, perhaps in the same school, and a small number of them might not enroll in college. We simulate the consequences of such responses in terms of the potential benefits to students in Section 8.

Understanding variation in performance across programs and implications for policy design

In this section, we discuss the patterns of low performance in our simulations with an eye towards addressing important policy concerns around the types of institutions, students, and fields of study that would be most affected by an accountability policy based on the net earnings premium and loan repayment rate metrics.

Poor performing programs or poor performing institutions?

Should performance measurement and corresponding sanctions primarily occur at the program or the institution level? One important consideration is the amount of variation in performance within institutions compared to the amount of variation between institutions. It is important to understand whether poor programs tend to be concentrated in institutions where performance is low across many or most programs—in which case institution level sanctions might be most appropriate—or whether they are spread across institutions which have a mix of good and bad programs. In the latter case, institution level metrics might miss poor performing programs and fail to protect the students enrolled in them and/or apply sanctions to high performing programs within an institution where many

programs fail. The concentration of poor performing programs within institutions also gives a sense for how institutions might be affected by accountability, by illustrating how much of their enrollment and thus revenue might be at risk.

Table 5 provides insight about how concentrated failing programs are within institutions by showing the distribution of institutions in each sector by the fraction of TIV students exiting from failing programs. Across all sectors, 81 percent of institutions have no failing programs at the undergraduate level and 64 percent have no failing graduate programs, while 90 and 86 percent of institutions have fewer than half of their undergraduate and graduate Title IV enrollment in failing programs, respectively. At the other end of the spectrum, 6 percent of institutions contain only failing undergraduate programs (among programs with reported data) while 9 percent have no passing graduate programs. Since the public sector represents the lion's share of enrollment, it is unsurprising that the distribution of enrollment in failing programs across public institutions is broadly similar to the overall distribution. However, only 2 and 3 percent of public institutions have all failing undergraduate and graduate programs, respectively. In the for-profit sector, we observe more of a bimodal distribution: 76 percent of institutions have no failing undergraduate programs, whereas 17 percent of institutions have no passing undergraduate programs. Almost two-thirds of for-profit institutions with graduate programs have at least one failing graduate program and 29 percent have no passing programs at the graduate level.[29]

Some sets of institutions warrant particular attention due to the access they provide to underserved populations. As shown in Table 6, over three-quarters of all Minority Serving Institutions (MSIs) had no failing undergraduate programs and, in 88% of these schools, fewer than a quarter of students attended failing programs. Among Historically Black Colleges and Universities (HBCUs), the median institution has no failing undergraduate programs, and 80 percent of institutions have less than a quarter of their undergraduate enrollment in failing programs. While no HBCUs have all undergraduate programs failing both metrics, about 7 percent have more than half of their undergraduate enrollment in failing programs. At the graduate level, many more HBCUs and MSIs have all or the majority of enrollment in failing programs. It is important to note, however, that we observe sufficient loan repayment and earnings information for only a small share of all HBCU

. . .

29. Single program institutions are more common in the for-profit sector and 90 percent of the institutions with 100 percent of students in failing programs have only one program with a nonmissing NEP and LRR. Note that 48 percent of these institutions offer more than one program, but the other programs were missing the necessary components to determine performance. Thus, it is likely that at least some of these institutions that show up as having only failing programs in our calculations would contain at least one passing program if earnings had been measured at the 2- versus 4-digit CIP code level.

graduate programs. For instance, among all HBCUs with only failing graduate programs, the average number of graduate program offerings is approximately 8 while the average number of programs with sufficient data to calculate loan repayment and earnings metrics is just over 1. Given the small number of HBCUs relative to the overall size of the higher education sector, these results may be less precise than estimated performance for larger subsets of institutions.[30]

Finally, we project that 68 percent of public community colleges will have no failing programs and 81 percent of community colleges will have fewer than a quarter of their students in failing programs.   Approximately 23 percent of failing programs in community colleges are associate degrees in general studies/liberal arts which serve 44 percent of community college students exiting failing programs.

In summary, programs that fail both the repayment and the net earnings metric tend to coexist in institutions where most students attend programs with positive net earnings, a positive repayment rate, or both. The for-profit sector is an exception, where among institutions with at least one failing program, 77 percent of the institutions have more than two-thirds of their enrollment in failing programs, and 67 percent have no passing programs. However, this is in part due to the higher prevalence of small institutions that have only one or relatively few different program offerings. The evidence here thus underscores the importance of our proposed program-based approach to accountability: institution-based metrics and sanctions may be too coarse to identify and protect students from poorly performing programs.

Performance and student demographics

A second key question raised by accountability systems relates to how measured performance varies with the characteristics of students served by different programs. A particular concern is whether programs that serve students who are more disadvantaged in terms of their likely labor market earnings or ability to repay their loans will be unfairly punished, cutting off access to higher education for students. This is a critically important issue, and there is no doubt that student characteristics like race, ethnicity, gender, family wealth, and prior academic preparation influence outcomes after college. However, our focus is not on measuring program 'quality' and ranking programs from best to worst, but in establishing

. . .

30.   Only 74 out of 101 HBCUs have at least one undergraduate program with nonmissing NEP and LRR metrics and only 37 have at least one graduate program that meets this criteria. Among the 687 MSIs, 525 have a least one undergraduate program and 220 have at least one graduate program with nonmissing NEP and LRR metrics.

a minimally acceptable level of performance for programs to warrant support from federal student aid: again, the standard we propose is that programs do not make students worse off financially. In this light, adjusting our metrics or thresholds for student demographics would perversely permit more financial harm to students who need protection the most. Moreover, one problem we wish to address is the history of some low-quality programs aggressively recruiting disadvantaged students and lowering standards for such programs would make this behavior easier.[31] At the same time, we should verify that our minimum standard is reasonable, in the sense that it does not set an unrealistic bar for programs that do serve high shares of disadvantaged students, or those likely to face subsequent discrimination in the labor market.

Unfortunately, due to data constraints, there is little information on student demographic characteristics available at the program level to accurately characterize their relationship with program outcomes. We instead categorize institutions based on the share of TIV exiters in failing programs and, using IPEDs data, summarize how demographics and other characteristics differ along this dimension.[32] Table 7 shows differences in institution-wide enrollment, whether undergraduate students received federal aid, and undergraduate students' demographic characteristics. Institutions where a greater share of TIV undergraduate students attend failing programs do have higher shares of Pell Grant recipients, although the differences are not dramatic: about 67 percent of students in institutions with no failing undergraduate students receive Pell grants, compared to about 87 percent of students receiving Pell at institutions with over 25 percent of their enrollment in failing programs. Black students are also overrepresented, in schools containing more students attending failing programs, while white students are under-represented, and this is especially true at the small number of institutions where all measured programs fail. There are similar, albeit slightly more muted, differences in the average age of students and the share who are first generation (i.e., neither parent complete college). Since IPEDs data report detailed demographic characteristics only for undergraduate students, we cannot similarly summarize graduate student characteristics by the share of graduate enrollment in failing programs.

. . .

31. See, for example, GAO (2010) and Cottom (2017).

32. The resulting sample is limited to the 3858 institutions that had at least one program with net earnings and loan repayment rate metrics (if the school participates in federal loan programs) or at least one program with net earnings (if the school does not participate in federal loan programs). Of these institutions, 173 had only graduate programs with nonmissing accountability metrics. 1663 institutions are excluded from these analyses because they did not have any programs with nonmissing accountability metrics. Appendix Table A9 contains characteristics of these excluded schools.

Using these same categories, Table 8 shows the breakdown of higher education institutions by sector, where sector is defined by control and predominant degree. Notably, 81 percent of schools where all undergraduate students attend failing programs are for-profit certificate-granting institutions (whereas only 27 percent of all institutions are in this sector). In contrast, public certificate and associate degree granting institutions as well as HBCUs and MSIs are underrepresented in the set of institutions where all programs fail. Turning to the distribution of schools across sectors according to graduate program performance, most schools with only failing programs are in the public and nonprofit sector but this is due to the relatively small number of for-profit schools with graduate program offerings.

Finally, Table 9 shows differences in institutional characteristics, including other accountability metrics such as cohort default rates and the borrower-based repayment rates from the College Scorecard. We again focus on characteristics according to the percent of undergraduate students in failing programs because almost all of these characteristics are specific to undergraduate students.[33] Many of the measures of institution-wide outcomes are correlated with the percent of undergraduate students in failing programs. The share of undergraduate students earning more than $25,000 ten years after first entering college and the share of undergraduate students paying down $1 of debt within 3 and 7 years of exiting fall as the share of students in failing programs increases. Interestingly, student loan cohort default rates – the current metric used to determine Title IV eligibility – is less correlated with the share of students in failing programs, although default rates are lower for schools with no failing undergraduate programs (9 percent versus 14-16 percent for schools with at least 1 failing program).

The bottom of Table 9 shows how program-level and school-level performance might interact.[34] Thirteen percent of schools where all students attending passing programs have negative loan repayment rates while 95 percent of schools where all students attend failing programs would fail the repayment rate metric if it were measured at the institution-level.

. . .

33. An exception is that the institutional finance items (tuition per FTE and instructional expenditures per FTE) include revenue and expenditures from both undergraduate and graduate programs. Likewise, institutional accountability metrics include both graduate and undergraduate programs.

34. The institution-level LRR is computed as the sum of balances at year 3 divided by the sum of balances at repayment entry across all programs at the institution (in other words, the methodology is the same as the program level LRR, except that balances from programs too small to be reported individually are also included). The institution-level NEP is computed by taking the weighted average of program NEPs within an institution but does not include any additional information on students in smaller programs where earnings data are not reported.

We estimate that 23 percent of schools with no enrollment in failing undergraduate programs would fail an institution-level NEP metric that combines undergraduate and graduate performance.

While Table 7 makes the point that student demographics are related to measured performance, the data clearly show that demographics are not destiny; or that programs that serve disadvantaged students can indeed meet the minimum standards we've outlined in this paper. For example, among schools that are in the highest quartiles of both Pell Grant receipt and the share of the undergraduate student body that are under-represented minorities, 13 percent of programs (with 87,130 students exiting each year) have net earnings in the top quartile of the NEP metric for the program's credential level and 12 percent of programs (44,594 students) have loan repayment rates in the top quartile.

Performance in programs with high social value but low pay

Another common concern with a program-level outcomes-based accountability system is that programs whose graduates work in critical, but lower paid industries or occupations might be disproportionately affected.

We explore this for two often cited examples: education and social services/public administration. The majority of students pursuing sub-baccalaureate credentials in education attended failing programs. Specifically, 52 percent of undergraduate students seeking a certificate in an education-related field and 61 percent of education-related associate degree seeking students attended a program where most students didn't make progress repaying their loans or earn more than a high school graduate after accounting for their out of pocket costs.[35] At the baccalaureate level, however, we find a very different pattern. Only 1.5 percent of bachelor's degree seeking students pursing an education-related major attended failing programs. Graduate program performance varies, with 3.4 percent of graduate certificate-seeking education students, 12 percent of master's degree students, and 57 percent of first professional degree students in education fields attending programs that fail both metrics. However, these high failure rates are almost completely driven by graduate programs in for-profit institutions. No first professional or doctoral education programs and

. . .

35. Appendix Table A10 displays the share of students attending failing programs by credential level and field (2-digit CIP code) while Appendix Table A11 displays the share of students in failing programs by credential level, field, and institutional control.

very few graduate certificate education programs in public and non-profit institutions fail both metrics.

The other broad field where this might be a concern includes social services and public administration programs. At the associate degree level, we again see fairly high failure rates, with 32 percent of students in public institutions and 19 percent of students in non-profit institutions attending failing programs. Once again, very few students in this field attend bachelor's degree programs with negative earnings premia and repayment rates (1.3 and 2.6 percent for public and nonprofit bachelor's degree programs, respectively). No graduate certificate or first professional degree programs focusing on public administration and/or social services fail both metrics but 20 percent of students pursuing master's degrees in public institutions attend failing programs, while 28 percent of nonprofit and for-profit master's degree students attend public administration/social services programs with negative loan repayment rates and net earnings premia. More research and better data are needed, but a tentative conclusion is that students in programs where credentials are closely tied to work in high social value careers do not seem particularly adversely impacted by our proposed metrics. Given the conservative earnings thresholds we use, we would expect programs with reasonable success finding stable employment for their graduates would not be at risk in our scheme.

Performance over the business cycle

A final aspect of program performance we consider is the extent to which local labor market conditions affect program performance. Recall that the NEP metric is based on the median earnings of all students that separate from the program in a given year and have positive earnings in the relevant reference year, minus an estimate of program costs and state-specific median earnings of high-school graduates (for undergraduate programs) or bachelor's degree recipients (for graduate programs). This way of constructing net earnings guards against adverse effects of local area conditions on measured performance in several ways. First, since only the earnings of employed workers are used to measure median program earnings, measured earnings for the program do not fall when workers exit employment entirely—rather, since workers exiting employment during economic downturns tend to have lower earnings (Farber 2011), an increase in non-employment could increase median earnings of the program. More generally, if workers who experience reductions in work hours during a recession tend to have earnings below the earnings of the median worker in a program, then measured performance would not be affected.

Still, we build adjustments into our threshold level of wages to further hold programs harmless for local labor market conditions and the business cycle. First, we use state-specific measures of reference group median earnings to measure net earnings premia (Appendix Figures A5 and A6), reflecting the fact that the opportunity costs of attending college are different in different areas and to avoid holding programs in low wage states to too high a standard. In the case of graduate programs, we also allow reference group earnings to vary across broad field categories to account to account for the fact that the counterfactual earnings levels for students entering graduate programs may differ substantially depending on the demand for workers with similar credentials.

Nationally, median annual earnings of high school degree or GED holders is $19,400. But the lowest wage states, including Mississippi, West Virginia, Alabama, New Mexico, and Louisiana, all have median earnings for the same group at or below $16,025. On the other hand, some states—North Dakota, Wyoming, Minnesota, Hawaii, and New Hampshire— have higher median high school earnings and are held to a higher bar.

There is even more variation in median earnings for bachelor's degree holders across states, ranging from less than $33,000 in New Mexico, Idaho, and Mississippi and greater than $45,000 in New York, Connecticut, New Jersey, Maryland, and Massachusetts. Taking into account the field of an individual's bachelor's degree yields even more variation, although the ordering of median earnings by broad field appears to be largely consistent within states, with arts and humanities majors receiving the lowest median earnings and STEM and allied health majors receiving the highest (Appendix Figure A7).[36]

Across all states, the percent of undergraduate students in failing programs would grow by less than 1 percentage point – from approximately 7.1 percent to 7.8 percent – if a national high school earnings threshold was used in place of state-specific thresholds (Appendix Table A12). This small change at the national level does mask some large changes in the share of students who attend programs deemed to be failing in specific states (Appendix Figures A12 and A13). Not surprisingly, this group includes states with some of the lowest and highest median earnings of high school graduates. For example, the percent of undergraduate students in failing programs in Mississippi would increase from approximately 9 percent to almost 33 percent if the national high school earnings threshold was used. Likewise, Louisiana would go from around 5 percent to over 20 percent of students attending

. . .

36.   We do not find much evidence that programs in states with higher reference median earnings have more enrollment in failing programs, overall, and within specific sectors (Appendix Figures A8 through A11).

failing programs. At the other end of the distribution, using a state-specific threshold results in 18 percent of undergraduate students attending failing programs in Hawaii while using a national threshold lowers this share to 3 percent.

Turning to graduate programs, we examine the impact of moving from state and field-specific reference earnings to national median earnings for bachelor's degree holders. Using the national median would increase the share of graduate students in failing programs from 32 to 36 percent. Many states would see at least a 10 percent increase in the share of graduate students in failing programs.[37] A smaller number of states would see as dramatic changes in the opposite direction, although Nevada would go from having 62 percent of students in failing programs to no students in failing programs with this change. Most of the increases in failure rates would occur in the public and nonprofit sectors while decreases would predominantly be among for-profit graduate programs (Appendix Figure A15).

Since program level data from the College Scorecard is only available for two cohorts at most, it is not possible to evaluate how the fraction of programs that fail our metrics might evolve over the business cycle. Appendix Figure A16 shows the trend in measures of median earnings by degree level, using the same data (the ACS) and methodology we use to calculate our counterfactual earnings in the NEP metric. Median earnings measures are cyclical and tend to fall when unemployment rises and vice versa. Importantly, however, the sensitivity of earnings to the business cycle is smaller at higher degree levels. This builds in a further adjustment to changes over the business cycle, since the counterfactual earnings we use will automatically adjust downward by more than the anticipated fall in earnings for both undergraduate and graduate programs. If anything, our measure might "over-adjust" during downturns and result in overly lenient thresholds in such years.

Potential benefits of accountability / costs of the status quo

Estimating the costs and benefits of the accountability scheme outline here is difficult, both due to data constraints and a thin research base to guide assumptions about how institutions and students might respond to the information and incentives generated by such a

. . .

37.   These states include Alabama, Florida, Kentucky, Mississippi, New Hampshire, New Mexico, North Carolina, Oklahoma, South Carolina, and West Virginia.

policy. Still, a sense for the magnitude of benefits that such a system could yield is important to be able to think about the relative importance of this policy relative to others.

The main benefit of an accountability policy based on these metrics will likely come from shifting students from programs that produce poor outcomes to those that produce higher earnings and loan repayment rates. If programs that fail both metrics are deemed ineligible to use federal student aid, we expect that such programs will close and/or students will choose to pursue other options where they can use their TIV funds. Since sanctions apply at the program level, a likely result of program-level loss of eligibility would be that students would switch their enrollment to another passing program within the same institution.

For the purpose of estimating how much students in failing programs could gain by switching to a better program, we assume that passing undergraduate programs in the same institution serve as substitutes—and thus programs students will likely switch to if their program is deemed ineligible for federal student aid—for failing undergraduate programs, and likewise for graduate programs. For example, a student could switch from a failing associate degree program to a passing certificate program, but not to a master's degree program. In some institutions, however, there might not be passing programs for students to switch to. For students attending such programs, we first make the conservative assumption that they will not enroll in any higher education program and thus only benefit by not paying for a program in excess of what it is worth. However, based on student enrollment responses to previous accountability measures (e.g., Cellini et al. 2020), it is likely that many of these students would instead enroll in a similar program at a nearby institution. Thus, in supplemental analyses, we present rough estimates of the aggregate benefits of accountability assuming that students who do not have an in-school alternative will instead choose a program at the same level in a nearby institution.

Table 10 focuses on failing programs in schools where there is at least one passing program at the same level that could serve as an alternative for students who would have otherwise attended failing programs. [38] Panel A focuses on undergraduate programs and shows that if undergraduate students in failing programs were to switch to passing alternatives in the

. . .

38. Note that our estimate of the number of programs without passing alternatives in the same institution is likely a conservatively high estimate given the large number of programs for which we do not have sufficient information to determine whether the program would have a negative net earnings premium or repayment rate. In institutions with at least one failing program and no passing alternatives, there are almost 10 additional programs at the same level with missing data, on average.

same institution, they would face similar out-of-pocket expenses but would receive approximately $10,000 more per year in net earnings, on average. Undergraduate students in failing programs have, on average, almost 10 alternative programs in their same institution that would pass at least one of the two metrics.[39] The difference in net earnings ($12,719) and number of alternative passing programs (7.2) is similar at the graduate level (Panel B).

In total, 370,643 undergraduate students leaving failing programs each year could have attended a similar program within their same institution that likely would leave them with higher earnings and/or the ability to keep up with their loan payments. Given the difference in net earnings between the programs they attended and the better alternatives, each undergraduate cohort loses at total of $4.71 billion in net earnings *each year* from enrolling in a failing program instead of a passing alternative. Assuming this absolute difference in earnings persists, each undergraduate cohort loses $75 billion (in present value) over 20 years.[40] At the graduate level, we estimate that 190,615 students that attend failing programs could have attended a passing alternative program in their same institution. Each graduate student cohort loses over $3.60 billion in net earnings *each year* from not pursuing a degree within the same institution in a passing alternative program, and about $68 billion over a 25-year working career. It is important to emphasize that these aggregate losses apply to each additional cohort that enrolls in programs that produce poor earnings and loan repayment outcomes.

Turning to students in failing programs that may not have an alternative passing program within their institution, we first note that these students would still benefit, on average, even if they did not attend college, albeit by a smaller amount than if they could move to a better program (Table 11).[41] These students who make up approximately 18 percent of all exiters from failing undergraduate programs (81,323 out of 451,966) and 13 percent of exiters from failing graduate programs (29,115 out of 219,729) in a given cohort. However, students without passing alternatives are not distributed evenly across sectors. Among programs in public institutions, less than 5 percent of undergraduate and graduate students in failing programs attend an institution without at least one passing alternative program at the same level (12,814 out of 284,292 undergraduates and 2,777 out of 58,367 graduate

. . .

39. For such students, there are also more than 19 additional undergraduate programs with missing data – a group that likely includes at least some additional passing programs.

40. Here we assume an interest rate of 2.5 percent to discount the flow of earnings to the 3rd year after exit.

41. These calculations abstract from any social and/or nonpecuniary benefits of college attendance.

students, respectively). In contrast, more than half of all students in failing for-profit undergraduate programs lack a passing alternative in their same institution (65,445 out of 116,670 students).

Panel A of Table 11 shows that undergraduate students in failing programs that lack a passing alternative in the same institution lose $307 million each year after accounting for their out-of-pocket costs and what they could have earned had they not attended the program. The cohort as a whole accumulates $952 million in student loan debt which, based on repayment over the first three years after leaving school, will continue to grow in volume. Each cohort of graduate program exiters loses $373 million each year and accumulates a total of $2.6 billion in student loan debt (Panel B).[42]

Overall, the absence of an accountability policy generates losses of net earnings on the order of $9 billion dollars per year, or about $155 billion over students' working lives, for each cohort exiting from higher education programs every year. This does not take into account that students who we fail to protect from poor programs also have more difficulty managing their debt, and as a result pay more in interest, late fees, and costs associated with default and adverse impacts on their credit history. Despite the fact that some students might choose not to attend college at all due to a lack of alternative programs in the same institution, this is still a feature of our proposal rather than a bug (assuming no large non-pecuniary benefits of college enrollment), since failing programs leave students worse off relative to typical high-school graduates. Finally, we have not accounted for the likelihood that a substantial share of students in failing programs in institutions without passing programs would likely switch to better performing programs in other programs, and so the costs of the status quo described here are conservative in that regard.

Further considerations for policy

Fully specifying an accountability policy based on the metrics we propose above is beyond the scope of this report. In this section, however, we offer brief comments on some other policy decisions that should addressed, based on our modelling and recent research findings.

. . .

42. Of the students without an alternative to their failing program within their own institution, 87 percent of undergraduates and 89 percent of graduate students are within 30 miles of an institution with at least one passing program at the same level (Appendix Table A14).

Institution level metrics and "backstops"

While we argue that an accountability system should be built with a focus on programs, institution level measures and accountability should be included in a complete policy for several reasons. Most importantly, some programs within institutions may avoid program level accountability either due to low enrollment preventing the disclosure of accountability metrics, or because we may explicitly exempt programs to avoid undesirable consequences as discussed below. Even at the 2-digit CIP level, some programs have too few exiters per year to report earnings measures and repayment rates. "Rolling-up" exit cohorts from multiple years may be one workaround this issue. But to ensure as many students are protected as possible, we suggest aggregate institution-wide NEP and LRR metrics be calculated and that institutions be required to pass these metrics as a "backstop" to the program level accountability measures we propose.

Aside from ensuring more students are covered by accountability, institution level accountability serves a further role. In institutions where the majority of the overall student body is being harmed financially, students in failing programs are likely best served by choosing a different institution rather than simply changing programs at the same institution. While we expect such a policy to affect a small number of institutions, an institution-level backstop would help encourage institutions that fail the majority of their students to exit.

Coverage and exclusions

The results of our simulated performance metrics show clearly that sizeable numbers of students are enrolled in programs that expose them to financial harm across nearly all sectors of higher education. While low performance is concentrated in the for-profit sector, it is not unique to that sector. And graduate level programs, despite enrolling students expected to have more experience navigating higher education options, are not free of options that leave their students with labor market prospects too dim to justify the considerable expense involved for students. As a rule, the federal government should not support enrollment programs that leave students worse off and all programs should be covered by the accountability policy scheme we propose. To be clear, this is not an argument against providing protections where statutory authority permits, as in the 2014 Gainful Employment regulations that omitted degree programs outside the for-profit sector: protecting some students is better than protecting none. But the data above argue strongly for the need for policies that can protect students across all higher education sectors.

Should there be any exceptions to this rule? Perhaps. There may be practical considerations such that some types of programs end up unfairly judged, and tweaks to data collection procedures may be infeasible or ineffective in addressing the issues. One example is the CIP code for "General Studies" which may be a 'catch-all' category for students who have yet to declare a major, and so might reflect outcomes for a disproportionate share of students who fail to complete their degrees. With bachelor's degree programs, it may not be desirable to pressure institutions or students to declare a major earlier in their career and including general studies in accountability would likely create such pressure.

For sub-baccalaureate programs the issue is more complex. Research has found that failure to support their students in finding a pathway towards a program of study quickly is a key reason for low completion rates (Jenkins and Cho 2012), and accountability pressure could help encourage progress towards such reforms. At the same time, students in Liberal Arts/General Studies programs in community colleges are often aiming to transfer to four year schools and deeming such programs ineligible for federal aid could harm the ability of community colleges to support transfers—a crucial pathway to socioeconomic mobility for many students.

How to account for non-completers is a key issue with any type of program-level accountability policy, and the solution might require computing the program-level metrics only, for example, for students who declare a major or accumulate half the required credits towards a degree. In administrative data, for example, ED could estimate program level earnings measures using cohorts of students that are enrolled for at least one or two full-time equivalent semesters. In such cases, institution level metrics can and should be used to capture the outcomes of all students including those who dropout prior to declaring a major or meeting the enrollment thresholds to be included in program outcome metrics. This should be less of an issue at different degree levels and outside of Liberal Arts/General Studies, where declaring a major at or near entry is the norm (e.g., undergraduate certificate and first professional degree programs).

A common concern is whether accountability based on earnings outcomes might have undesirable impacts on programs whose graduates serve the public good, e.g. by working in early childhood education or social work, where earnings are generally low. First, we emphasize that the thresholds for the metrics described above represent low levels of earnings—for example, for undergraduate programs, threshold earnings (of HS graduates) are uniformly below the federal poverty guideline for a family of 4, and typical earnings for

workers with post-secondary credentials in these fields are comfortably above these levels.[43]

Nonetheless, concern about adverse effects might be further allayed, for example, by integrating our performance metrics with data collected through Public Sector Loan Forgiveness programs. It could be possible, for example, under a system that forgives some portion of student loans for each year a student is employed in the public sector to count such forgiveness as "repayment" to increase the likelihood that a program that produces students who work in the public sector will pass. To the extent possible, exceptions to the accountability metrics we outline should be based on student outcomes, as opposed to ex ante categorization of programs.

The last issue we address is how and whether to include institutions that serve communities that, and that as institutions themselves, have faced a legacy of discrimination that limits their students' financial success. Historically Black Colleges and Universities (HBCUs) have and continue to play critical roles in providing access to education for their communities. An accountability system that leads to disproportionate impacts on HBCUs would likely generate negative spillovers beyond those created by TIV loss in other sectors. We first note that at the undergraduate level, the majority of HBCUs have no undergraduate enrollment in programs that are predicted to fail both metrics and no HBCUs have all of their undergraduate programs fail (Table 6). At the graduate level, however, a larger share of programs are predicted to not pass the NEP and LRR metrics. For many of these programs, there may not be a substitute that would allow the student to remain in an HBCU. Although outside the scope of this proposal, we would advocate that accountability in this sector be paired with large-scale investments in HBCUs, of sufficient size to compensate for past discrimination and lack of access to funding, that would allow for reductions in student costs in the form of lower tuition and higher need-based institutional grants, and meet long standing infrastructure and other needs.[44]

. . .

43. There are only three states where the 25th percentile of earnings for bachelor's degree recipients with any annual earnings is less than median earnings of high school graduates (Idaho, Utah, and Wyoming) and on average, the difference between the 25th percentile of bachelor's degree recipient earnings and median high school earnings is $7690 nationwide.

44. Similar concerns and potential solutions exist for Tribal Colleges and Universities, although these institutions are smaller and fewer in number than HBCUs, so we are unable to generate reliable estimates of performance on the LRR and NEP for this sector.

Sanctions and other consequences

Most importantly, we recommend any accountability policy include sanctions that ultimately lead to loss of Title IV student aid eligibility for failing programs. In the past, similar institution-level sanctions tied to cohort default rates were effective in redirecting student enrollment away from low quality institutions (Looney and Yannelis 2019; Cellini et al. 2020).

The basic sanction structure we propose is that programs should become lose Title IV eligibility if they fail both the NEP and LRR metrics, for a certain number of years. In addition to this program level accountability, a full policy should include two important features. First, there should be institution-level measures of performance linked to institutional sanctions. Such measures would serve two purposes: first, there are many programs that are likely to have too few students to be able to measure and disclose program level earnings. A backstop institution-level metric would aggregate performance across students in different programs and in doing so, provide students enrolled in these smaller programs some measure of protection. Second, institution level sanctions should be invoked when a substantial share—e.g. more than half—of an institution's enrollment overall (or measured across programs) is being harmed financially based on the metrics we outline. As we document here, in most cases, failing programs are in institutions where most students are not in such programs and so targeting such programs and encouraging the institution to redeploy its resources elsewhere is a sensible response. If poor performance characterizes the majority of an institution's offerings, however, then an institution level sanction is appropriate to prevent further harm to students.

A second important feature for an effective policy is that institutions be held accountable for the financial harm caused by poor performing programs, especially in the event those programs close. Looney and Yannelis (2019) document that a strikingly large share of student loan defaults are driven by schools and programs that enter and exit loan programs following relaxations and tightening of rules around participation in federal student aid programs. Accountability policies could mitigate the incentive for bad programs to enter by holding institutions and their ownership accountable for some share of loans that are not repaid, or Pell dollars that don't lead to higher earnings. The metrics discussed here could be used to trigger and scale such risk sharing.

To proactively keep students away from low performing programs, the federal government should also disclose the metrics generated by the accountability system—and where possible report metrics at a finer level of detail (e.g., where enrollment is high enough, reporting

earnings and repayment rate for 4-digit CIP programs)—to support and encourage students and families to choose programs that are more likely to raise earnings and not burden students with debt they cannot repay. Again, research has found that in some cases, disclosures can influence student choices (e.g., Steffel et al. 2020), though typically the impact of existing disclosure regimes has been modest and concentrated among more affluent students who may have more supports and resources to help leverage better information into better college choices (Hurwitz and Smith 2018).[45]

Disclosures could be more effective with steps taken to ensure they reach prospective students, for example by using the electronic FAFSA to present applicants with performance metrics for any institution or program they express interest in, or by requiring institutions to document that students have received and acknowledged information disclosures. While disclosures may enable all students to make better choices and enable better monitoring and oversight of program quality by other stakeholders, they should supplement rather than substitute a consequential accountability scheme.

Conclusion

In this report, we illustrate a framework that could be used to build an accountability system aimed at preventing students from enrolling in programs that would result in financial harm. The core of our proposed system is a commonsense criterion that programs that harm their students financially should not be subsidized through eligibility for federal student aid programs. We measure the financial outcomes of students in two complementary ways: whether their net earnings are greater than they would have been had they not enrolled in the program and whether they able to make *any* progress repaying the loans.

As we discuss in previous sections, data limitations make it necessary to model the program earnings measure we propose, creating uncertainty in our simulations of the number of programs and students that would be impacted by our proposal and in the expected gains from such a policy. Importantly, the proposed metrics can be easily estimated by the Department of Education based on existing administrative data or cross-agency data sharing agreements similar to those already used in the past to produce the College Scorecard data, or the Gainful Employment earnings measures. Any Administration interested in exploring

. . .

45. However, there is also evidence that even well-designed information disclosures and tools may have no effect on high school students' decisions, especially if insufficient efforts are made to make students aware of them (Blagg et al. 2017), or that disclosures may lead to unintended responses on the part of institutions (Baker 2020).

policies to protect students in the manner we've outlined here can and should replicate the analyses presented here with more accurate data.

We estimate that about 670,000 students per year—or 9 percent of all students that exit post-secondary programs on an annual basis—attended programs that leave them worse off financially. Under reasonable assumptions, for each cohort of students, the increase in earnings over their working career likely exceeds $155 billion. These numbers suggest that implementing a set of minimum requirements for programs to participate in federal aid programs accountability policy could be an important tool to protect students from substantial harm and improve postsecondary student outcomes.

References

Armona, L., Chakrabarti, R., & Lovenheim, M. (2020) "Student Loan Debt and Default: The Role of For-Profit Colleges," Federal Reserve Bank of New York Staff Report No. 811.

Baker, Dominique J. (2020) "'Name and Shame': An Effective Strategy for College Tuition Accountability?" *Educational Evaluation and Policy Analysis,* 42(3): 393-416.

Bettinger, E., & Soliz, A. (2016) "Returns to Vocational Credentials: Evidence from Ohio's Community and Technical Colleges," CAPSEE Working Paper, Center for Analysis of Postsecondary Education and Employment.

Bettinger, Eric P., Lindsay Fox, Susanna Loeb, & Eric S. Taylor (2017) "Virtual classrooms: How online college courses affect student success," *American Economic Review* 107(9): 2855-2875.

Blagg, K., Chingos, M. M., Graves, C., Nicotera, A., & Shaw, L. (2017). Rethinking consumer information in higher education. Urban Institute Research Report. Available at: https://www.urban.org/sites/default/files/publication/91666/rethinking_consumer_information_in_higher_education_2.pdf.

Brickley, James A. & R. Lawrence Van Horn (2002) "Managerial Incentives in Nonprofit Organizations: Evidence from Hospitals," *Journal of Law and Economics*, 45(1): 227-249.

Capps, Cory S., Dennis W. Carlton, & Guy David (2020) "Antitrust Treatment of Nonprofits: Should Hospitals Receive Special Care?" *Economic Inquiry*, 58(3): 1183-1199.

Cellini, S. Riegg and L. Chaudhary (2014) "The Labor Market Returns to a For-Profit College Education", *Economics of Education Review*. 43: 125-140.

Cellini, Stephanie R., Rajeev Darolia, and Lesley J. Turner (2020) "Where Do Students Go When For-Profit Colleges Lose Federal Aid?" *American Economic Journal: Economic Policy*, 12 (2): 46-83.

Cellini, S. R., & Turner, N. (2019) "Gainfully employed? Assessing the employment and earnings of for-profit college students using administrative data," *Journal of Human Resources*, 54(2): 342-370.

Chetty, Raj, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. (2017) "Mobility Report Cards: The Role of Colleges in Intergenerational Mobility." *Opportunity Insights Working Paper.*

Chou, Tiffany, Adam Looney, and Tara Watson (2017) "Measuring Loan Outcomes at Postsecondary Institutions: Cohort Repayment Rates as an Indicator of Student Success and Institutional Accountability," NBER working paper 23118.

Cottom, Tressie McMillan (2017) *Lower Ed: The Troubling Rise of For-Profit Colleges in the New Economy*. New York: The New Press.

Darolia, R. (2013) "Integrity versus access? The effect of federal financial aid availability on postsecondary enrollment," Journal of Public Economics, 106: 101-114.

Dunlop, Erin (2013) "What do Stafford Loans Actually Buy You? The Effect of Stafford Loan Access on Community College Students," CALDER Working Paper 94.

Duggan, Mark (2000) "Hospital Ownership and Public Medical Spending," *Quarterly Journal of Economics*, 115(4): 1343-1373.

Farber, Henry A. (2011) "Job Loss In The Great Recession: Historical Perspective From The Displaced Workers Survey, 1984-2010." *National Bureau of Economic Research Working Paper* 17040.

Gaulke, A., Cassidy, H., & Namingit, S. (2019) "The effect of post-baccalaureate business certificates on job search: Results from a correspondence study," *Labour Economics*, 61: 101759.

Gaynor, Martin & William B. Vogt (2002) "Competition Among Hospitals," *Rand Journal of Economics*, 34(4): 764-785.

Government Accountability Office (2010) "For-Profit Colleges: Undercover Testing Finds Colleges Encouraged Fraud and Engaged in Deceptive and Questionable Marketing Practices." GAO-10-948T.

Government Accountability Office (2018) "Federal Student Loans: Actions Needed to Improve Oversight of Schools' Default Rates." GAO-18-163.

Hart, Cassandra M.D., Elizabeth Friedmann, & Michael Hill (2018) "Online Course-taking and Student Outcomes in California Community Colleges," *Education Finance and Policy* 13(1): 42-71.

Hoxby, Caroline M. (2019) "Online Postsecondary Education and Labor Productivity." In Charles R. Hulten and Valerie A. Ramey (Eds.), *Education, Skills, and Technical Change: Implications for Future US GDP Growth*. University of Chicago Press.

Hurwitz, Michael and Jonathan Smith (2018) "Student Responsiveness to Earnings Data in the College Scorecard," *Economic Inquiry* 56(2): 1220-1243.

Jenkins, Davis and Sung-Woo Cho (2012). *Get with the program: Accelerating community college students' entry into and completion of programs of study* (CCRC Working Paper No. 32). New York, NY: Community College Research Center, Teachers College, Columbia University.

Kelchen, Robert & Zhuoyao Liu (2019) "Did Gainful Employment Regulations Result in College and Program Closures? An Empirical Analysis," working paper. Available at: https://kelchenoneducation.files.wordpress.com/2019/11/kelchen_liu_nov19.pdf.

Liu, V. Y. T., C. R. Belfield, & M. J. Trimble (2015). "The medium-term labor market re-turns to community college awards: Evidence from North Carolina," *Economics of Education Review*, 44: 42–55.

Looney, A. & C. Yannelis (2019) "The consequences of student loan credit expansions: Evidence form three decades of default cycles," working paper.

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.

Silverman, Elaine & Jonathan Skinner (2004) "Medicare Upcoding and Hospital Owner-ship," *Journal of Health Economics*, 23(2): 369-389.

Steffel, Mary, Dennis Kramer II, Walter McHugh, and Nick Ducoff (2020) "Informational Disclosure and College Choice," Brookings Institution.

Turner, Lesley J. (2016) "The Returns to Higher Education for Marginal Students: Evidence from Colorado Welfare Recipients," *Economics of Education Review*, 51: 169-184.

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).

Appendix

Figures and Tables

Figure 1A: Undergraduate Certificate Program Performance by Metric
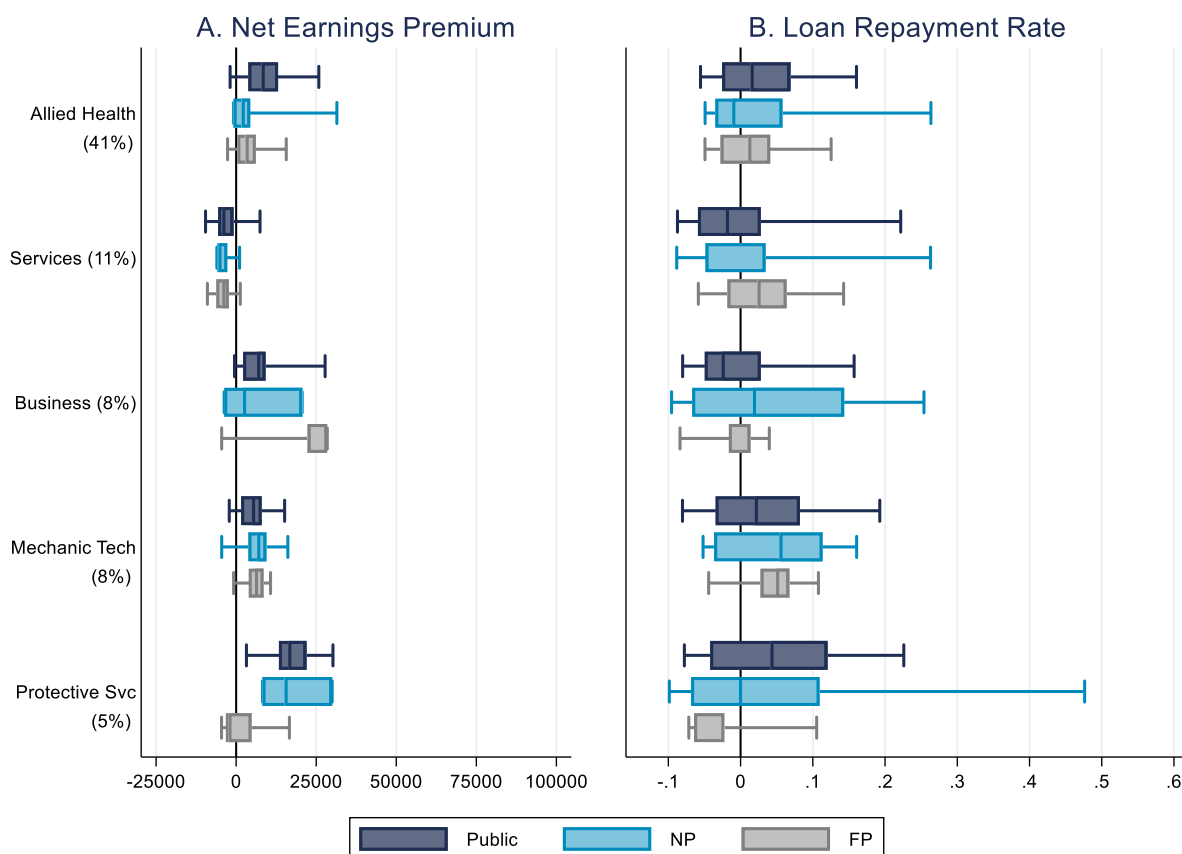
Notes: Top 5 fields (by % of total enrollment – shown in parentheses); see Appendix Figures A1 and A2 for all fields. Box represents interquartile range (25th to 75th percentile) of performance on metric, the vertical line in the center of the box indicates median performance, and the whiskers indicate performance at the 5th and 95th percentiles. Protective Services includes security programs. Mechanic Tech = mechanic and repair technologies/technicians. Services = personal and culinary services and family/consumer sciences.

Figure 1B: Associate Degree Program Performance by Metric
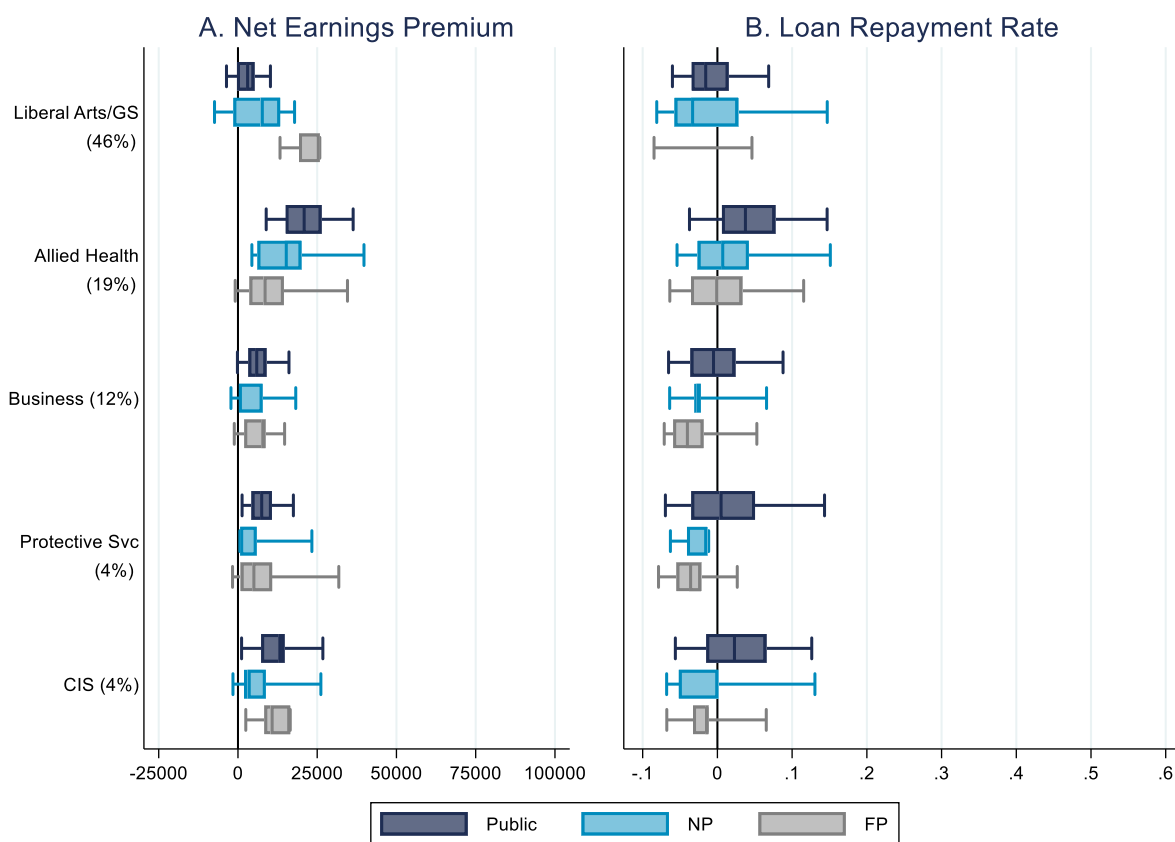
Notes: Top 5 fields (by % of total enrollment – shown in parentheses); see Appendix Figures A1 and A2 for all fields. Box represents interquartile range (25th to 75th percentile) of performance on metric, the vertical line in the center of the box indicates median performance, and the whiskers indicate performance at the 5th and 95th percentiles. CIS = computer and information sciences (including math and statistics). Protective Services includes security programs.

Figure 1C: Bachelor's Degree Program Performance by Metric

Notes: Top 5 fields (by % of total enrollment – shown in parentheses); see Appendix Figures A1 and A2 for all fields. Box represents interquartile range (25th to 75th percentile) of performance on metric, the vertical line in the center of the box indicates median performance, and the whiskers indicate performance at the 5th and 95th percentiles. CIS = computer and information sciences (including math and statistics). Social sciences includes psychology programs.

Figure 1D: Graduate Certificate Program Performance by Metric

A. Net Earnings Premium

B. Loan Repayment Rate

Notes: Top 5 fields (by % of total enrollment – shown in parentheses); see Appendix Figures A1 and A2 for all fields. Box represents interquartile range (25th to 75th percentile) of performance on metric, the vertical line in the center of the box indicates median performance, and the whiskers indicate performance at the 5th and 95th percentiles. Social sciences includes psychology programs. CIS = computer and information sciences (including math and statistics).

Figure 1E: Master's Degree Program Performance by Metric
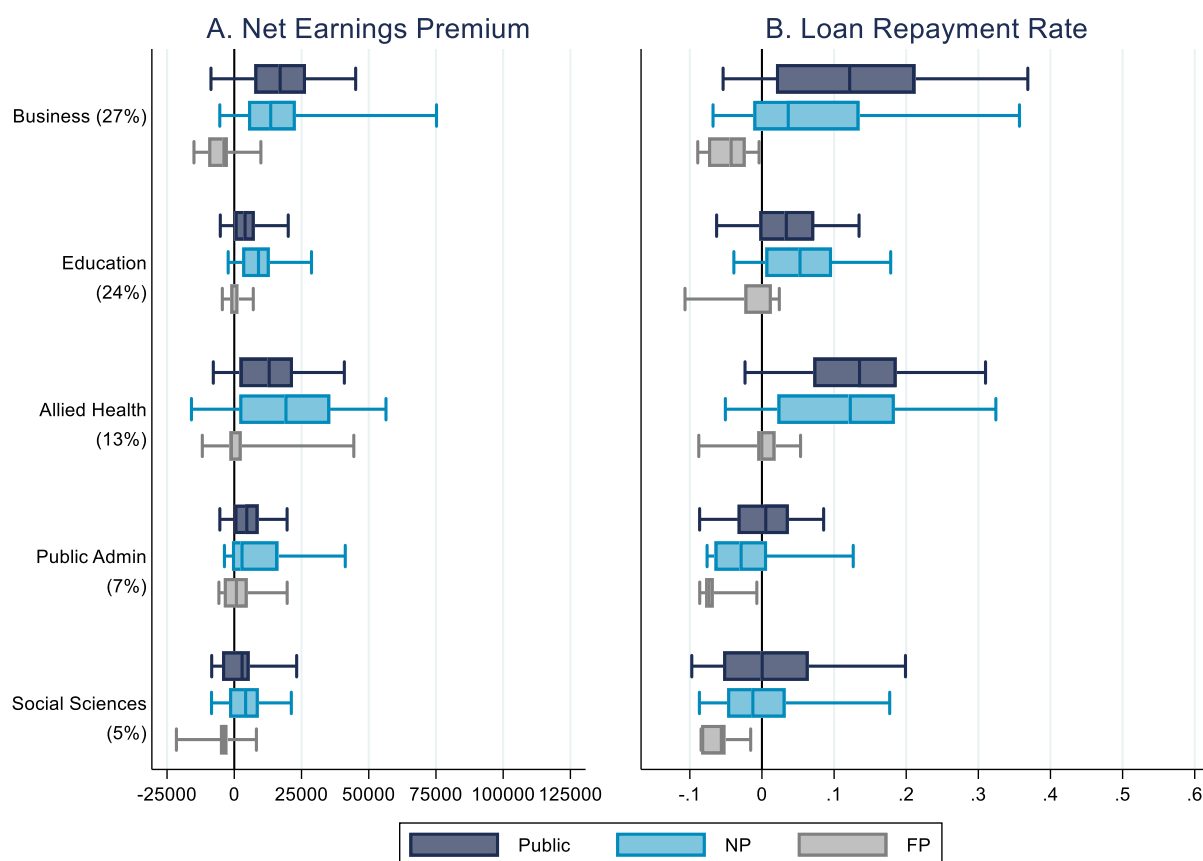
A. Net Earnings Premium   B. Loan Repayment Rate

Notes: Top 5 fields (by % of total enrollment – shown in parentheses); see Appendix Figures A1 and A2 for all fields. Box represents interquartile range (25th to 75th percentile) of performance on metric, the vertical line in the center of the box indicates median performance, and the whiskers indicate performance at the 5th and 95th percentiles. Social sciences includes psychology programs.

Figure 1F: Doctoral Degree Program Performance by Metric
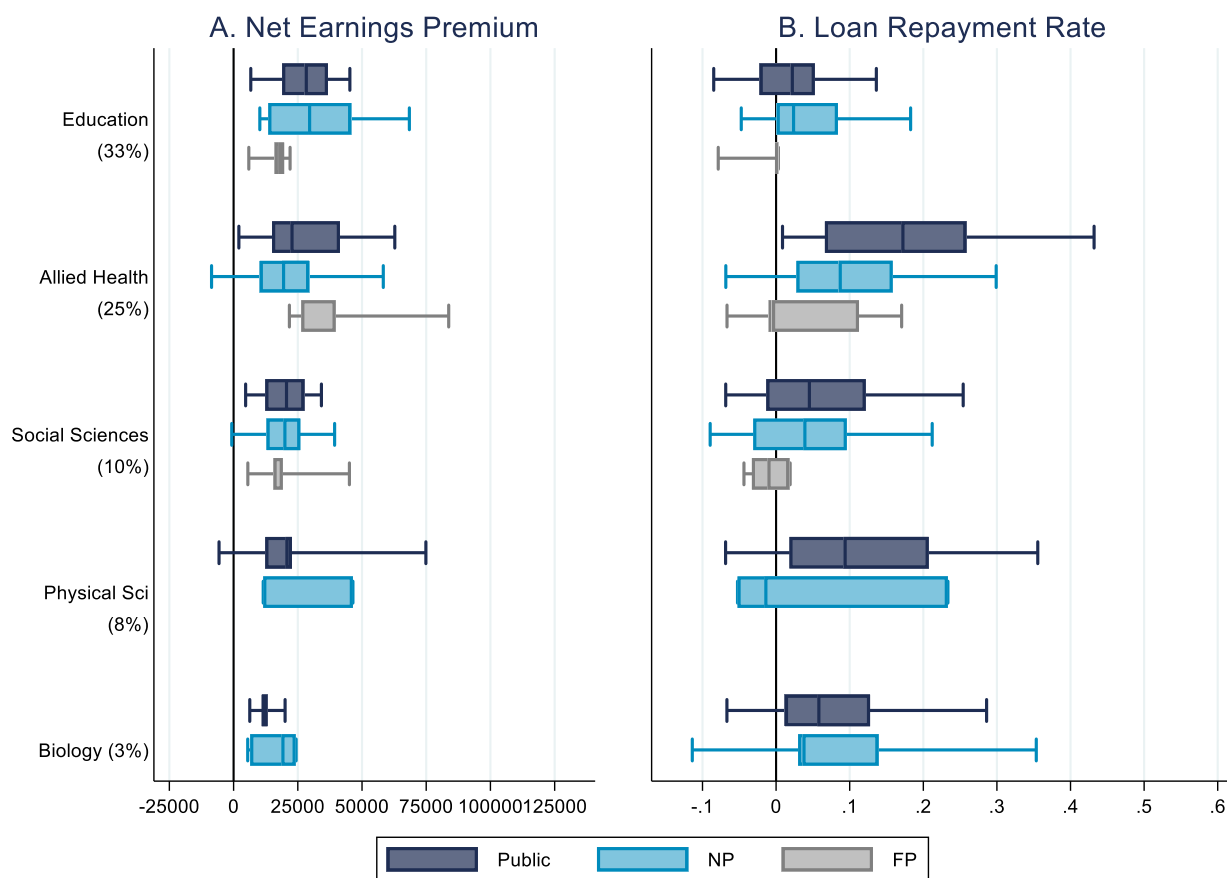
Notes: Top 5 fields (by % of total enrollment – shown in parentheses); see Appendix Figures A1 and A2 for all fields. Box represents interquartile range (25th to 75th percentile) of performance on metric, the vertical line in the center of the box indicates median performance, and the whiskers indicate performance at the 5th and 95th percentiles. Social sciences includes psychology programs.

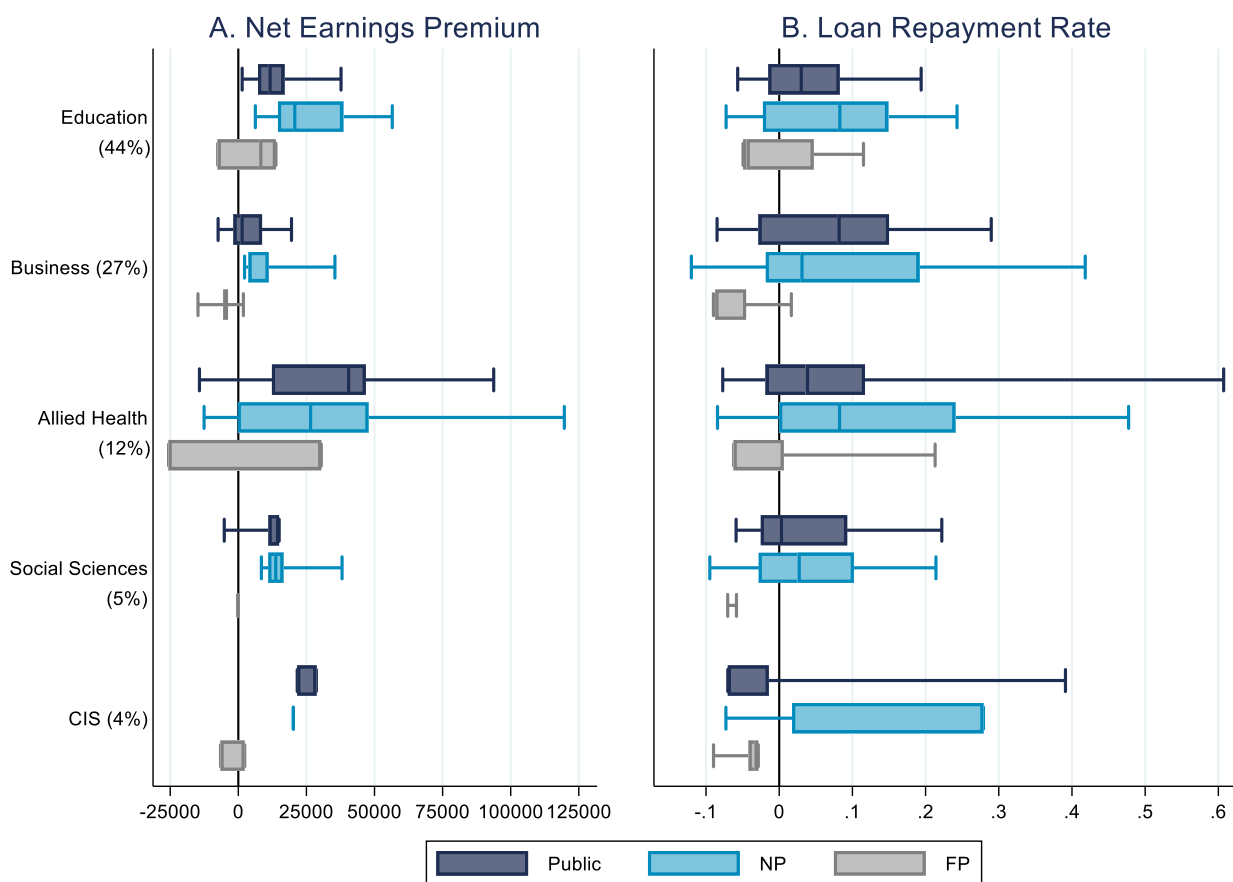Figure 1G: First Professional Degree Program Performance by Metric
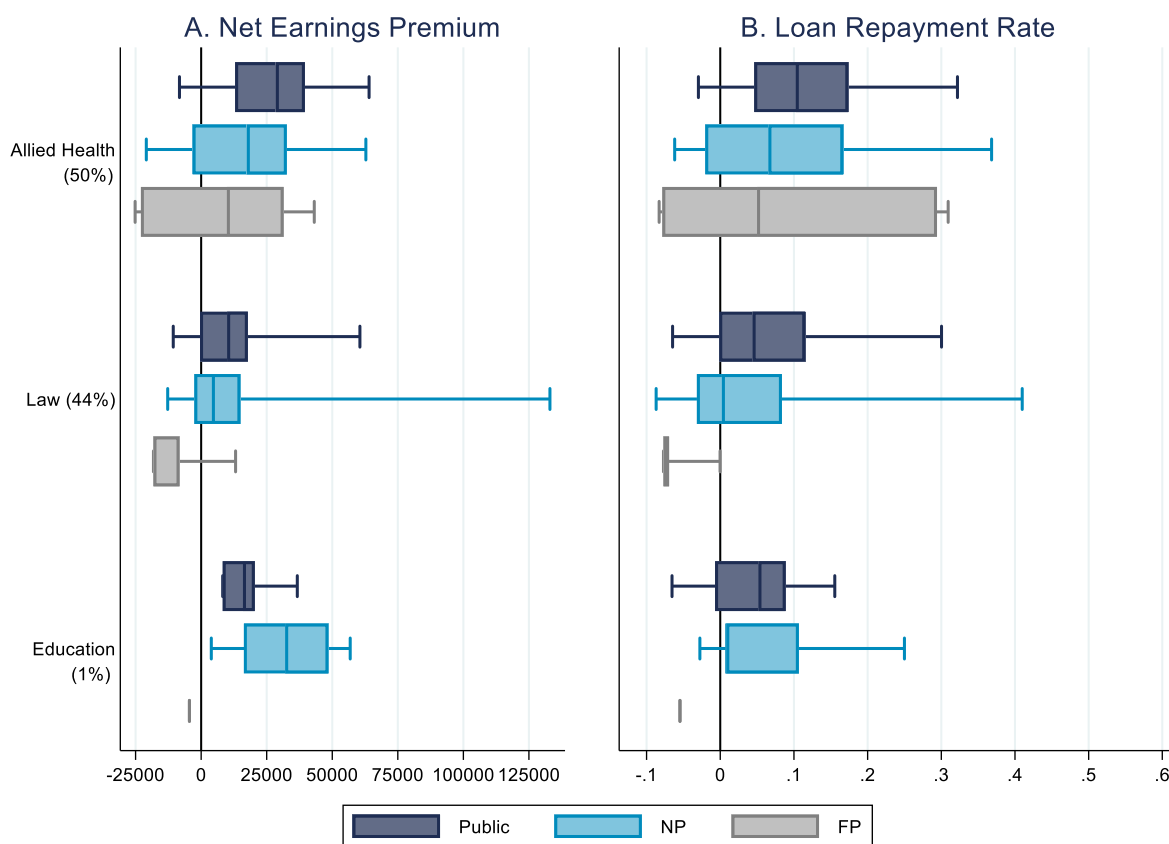
Notes: Top 3 fields (by % of total enrollment – shown in parentheses); see Appendix Figures A1 and A2 for all fields. Box represents interquartile range (25th to 75th percentile) of performance on metric, the vertical line in the center of the box indicates median performance, and the whiskers indicate performance at the 5th and 95th percentiles.

Figure 2A: Undergraduate certificate programs

A. Public Institutions

B. NP Institutions

C. FP Institutions

◇ Business	+ Mechanic/repair	☐ Security/protective svc
△ Personal svc	○ Allied Health	✕ All others
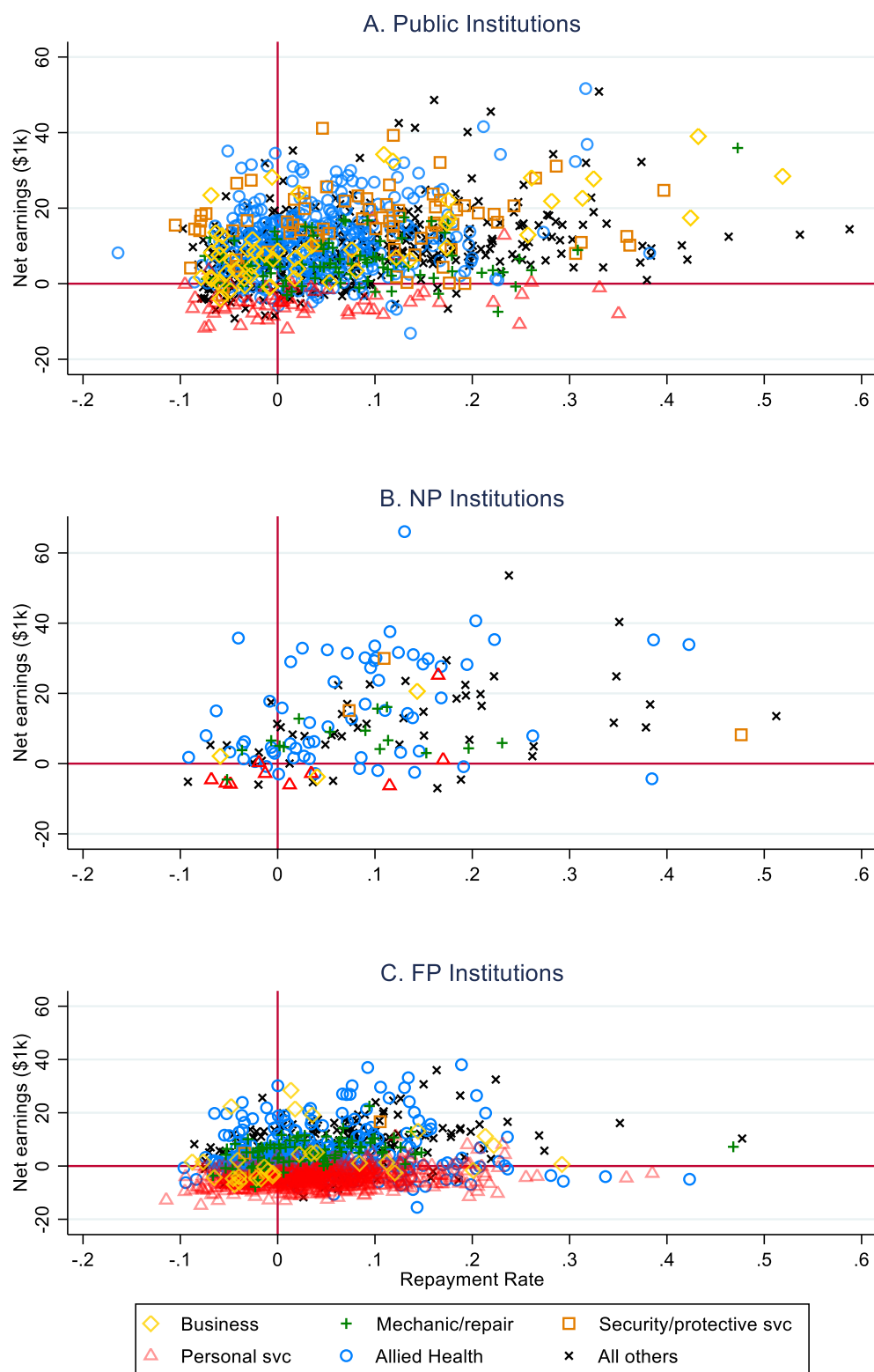
Figure 2B: Associate degree programs

A. Public Institutions

B. NP Institutions

C. FP Institutions

Business   CIS   Security/protective svc
Liberal Arts/GS   Allied Health   All others
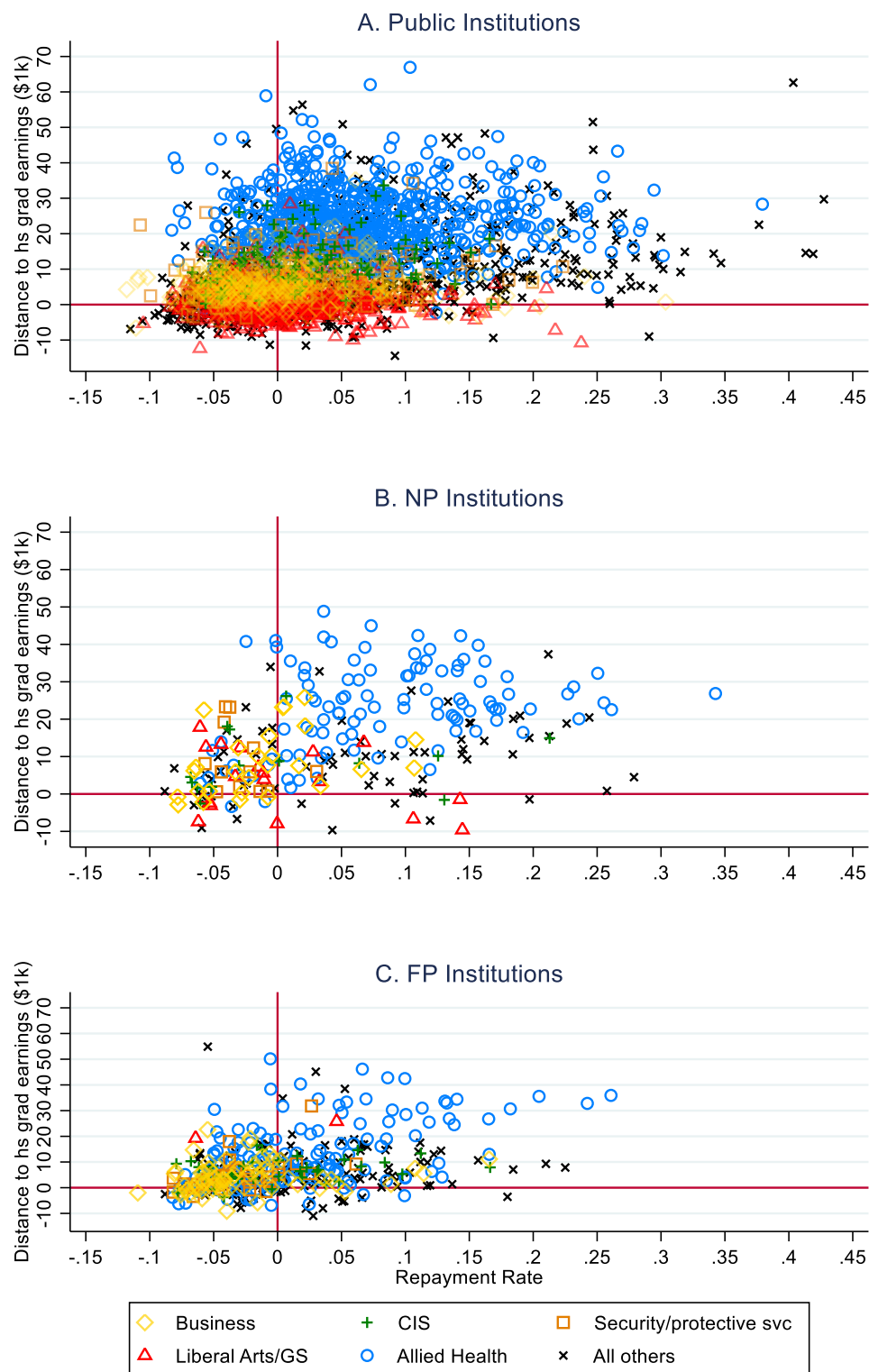
Figure 2C: Bachelor's degree programs

A. Public Institutions

B. NP Institutions

C. FP Institutions

	Business		CIS		Biology
	Social Science		Allied Health		All others
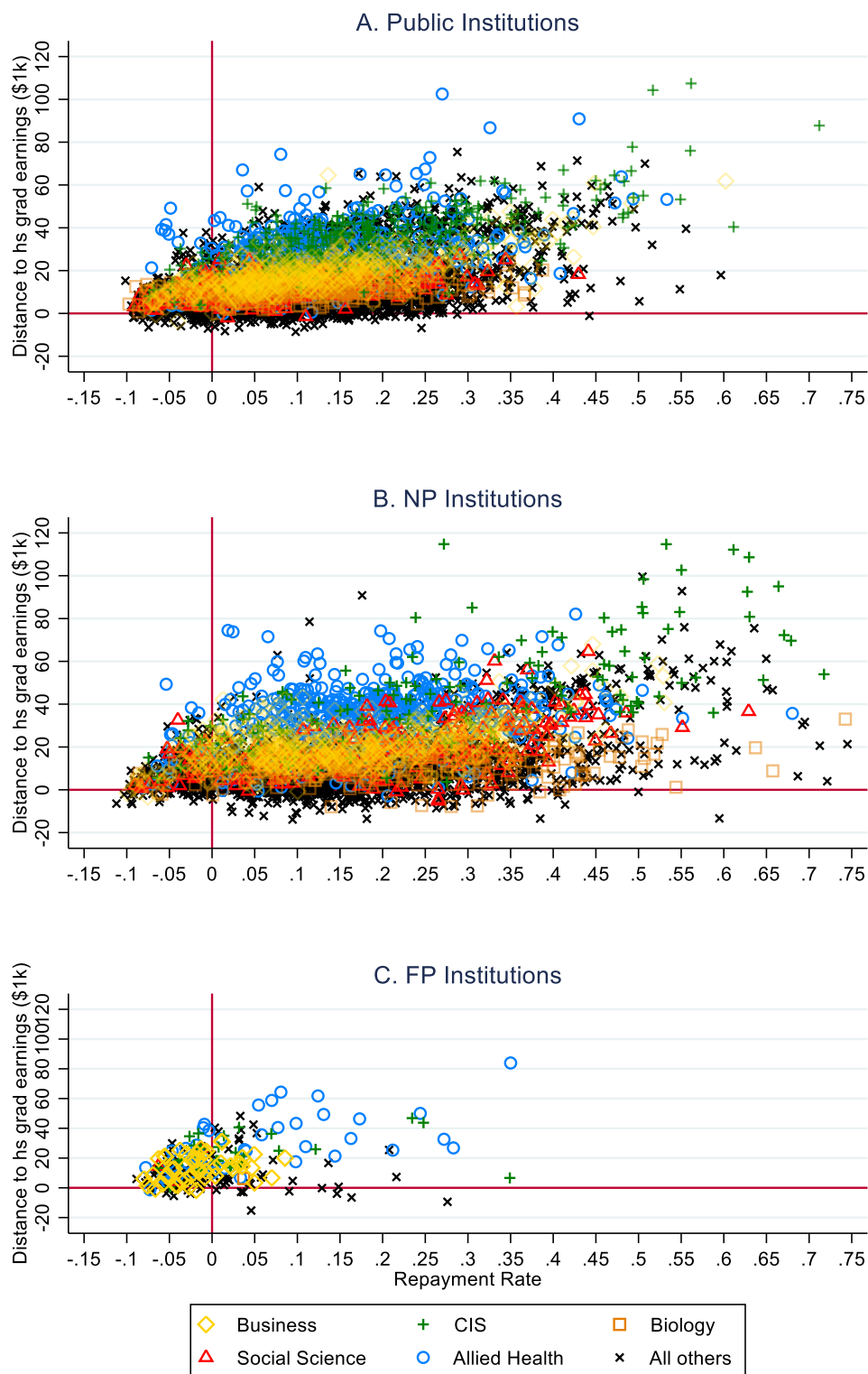
Figure 2D: Graduate Certificate Programs

A. Public Institutions

B. NP Institutions

C. FP Institutions

◇ Business	+ CIS	□ Social Sciences
△ Education	○ Allied Health	× All others
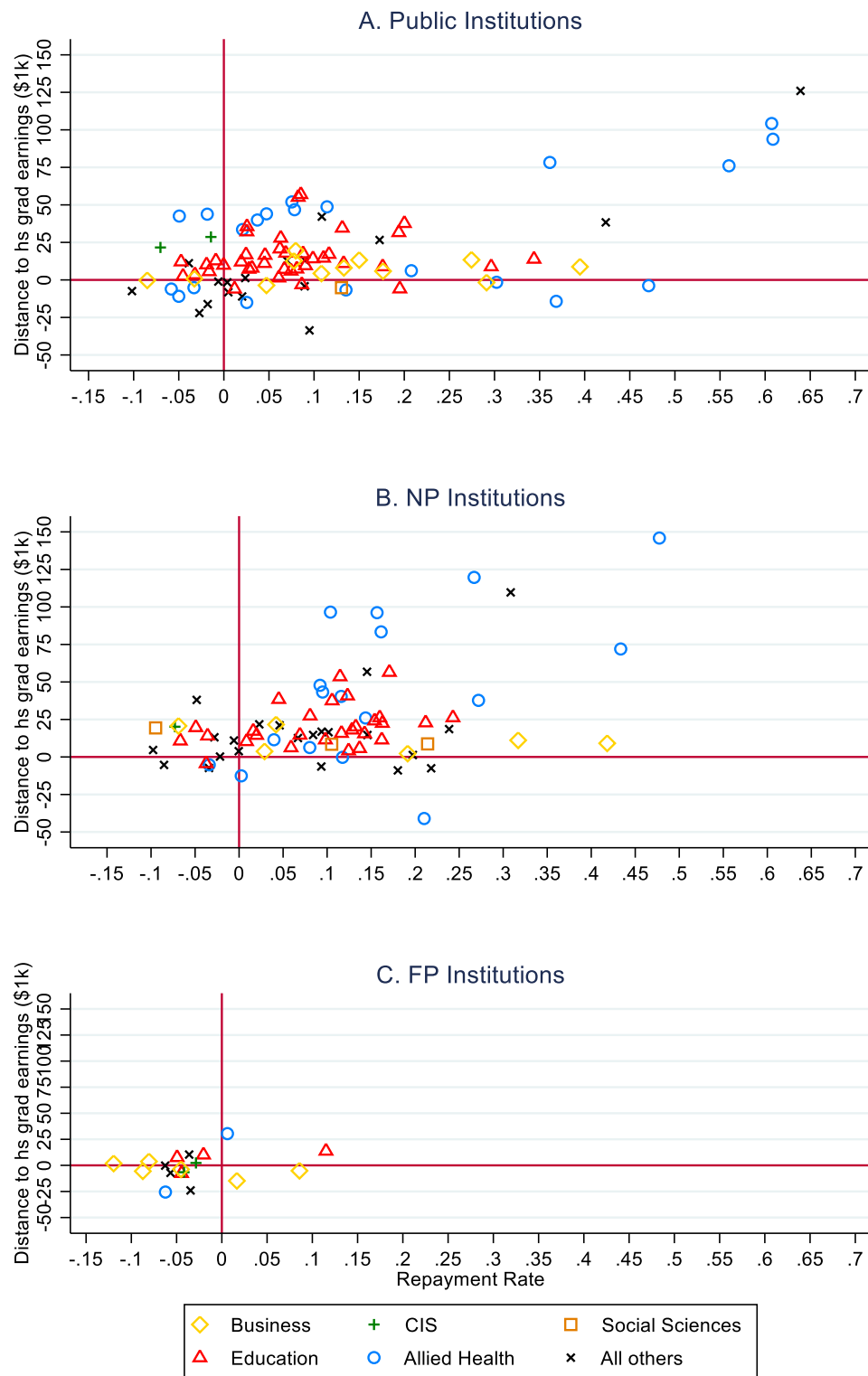
Figure 2E: Master's Degree Programs

A. Public Institutions

B. NP Institutions

C. FP Institutions

Legend:
◇ Business + Public Admin. □ Social Sciences
△ Education ○ Allied Health × All others
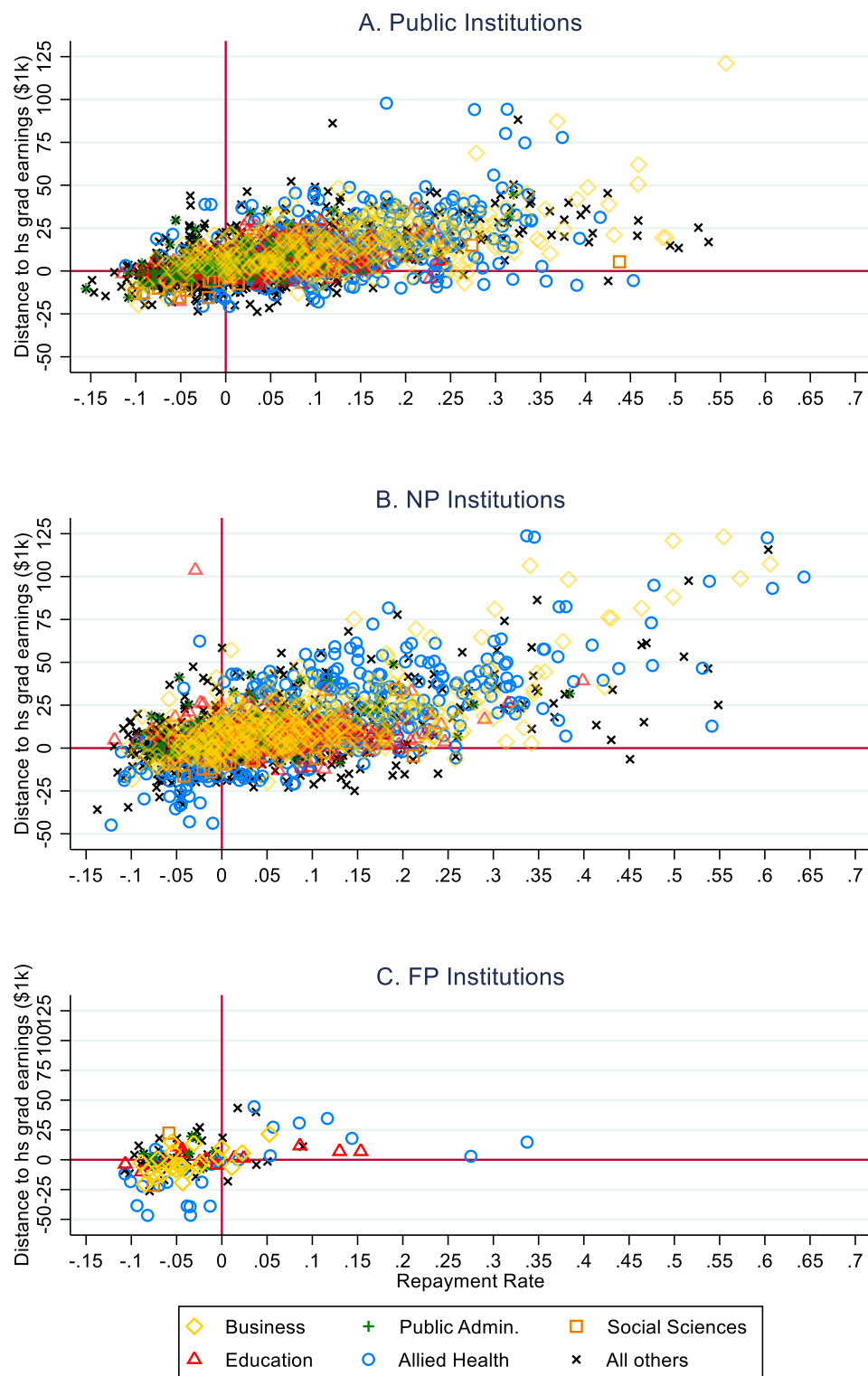
Figure 2F: Doctoral Degree Programs

A. Public Institutions

B. NP Institutions

C. FP Institutions

| ◇ | Business | + | Physical Sciences | □ | Social Sciences |
| △ | Education | ○ | Allied Health | × | All others |
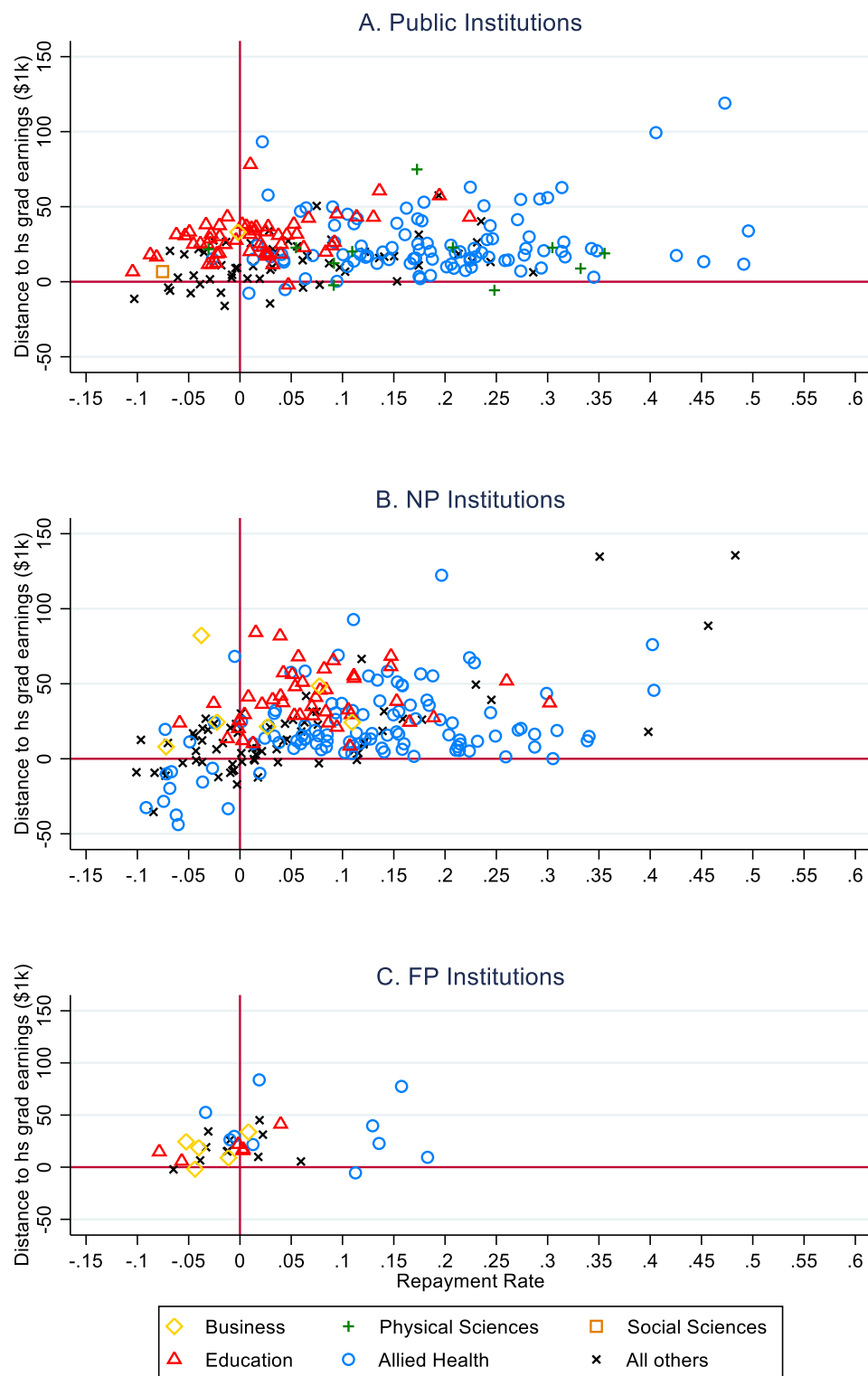
Figure 2G: First Professional Degree Programs

A. Public Institutions

B. NP Institutions

C. FP Institutions

◇ Business   + Law   △ Education   ○ Allied Health   × All others

Table 1: Programs with negative net earnings premia by credential level and institutional control

	(1) Undergrad. certificate	(2) Associate degree	(3) Bachelor's degree	(4) Grad. certificate	(5) Master's degree	(6) Doctoral degree	(7) First prof. degree	(8) All credentials
A. Reweighted to represent exiters from all programs								
All institutions	23%	17%	3%	25%	26%	7%	25%	15%
By control								
Public institutions	16%	18%	2%	19%	21%	9%	14%	12%
Nonprofit institutions	26%	13%	6%	11%	20%	7%	33%	14%
For-profit institutions	36%	12%	3%	72%	65%	3%	69%	30%
B. Unweighted								
All institutions	39%	18%	4%	23%	27%	10%	22%	15%
By control								
Public institutions	19%	18%	3%	24%	26%	7%	12%	11%
Nonprofit institutions	23%	12%	6%	14%	26%	14%	29%	13%
For-profit institutions	60%	19%	8%	56%	52%	9%	54%	45%

Notes: The net earnings premium metric represents the difference between median cohort earnings (net of median out-of-pocket costs) and median reference group earnings, where the reference group for undergraduate programs is high school graduates and the reference group for graduate programs is college graduates in the same broad field. See Sections 4 and 5 for additional details. Only programs with a nonmissing net earnings premium (N = 22,669) are included. Panel A estimates are weighted by the number of estimated exiters from a given program (see Section 5 and Appendix B for details). Panel B is unweighted.

Table 2: Programs with negative loan repayment rates by credential level and institutional control

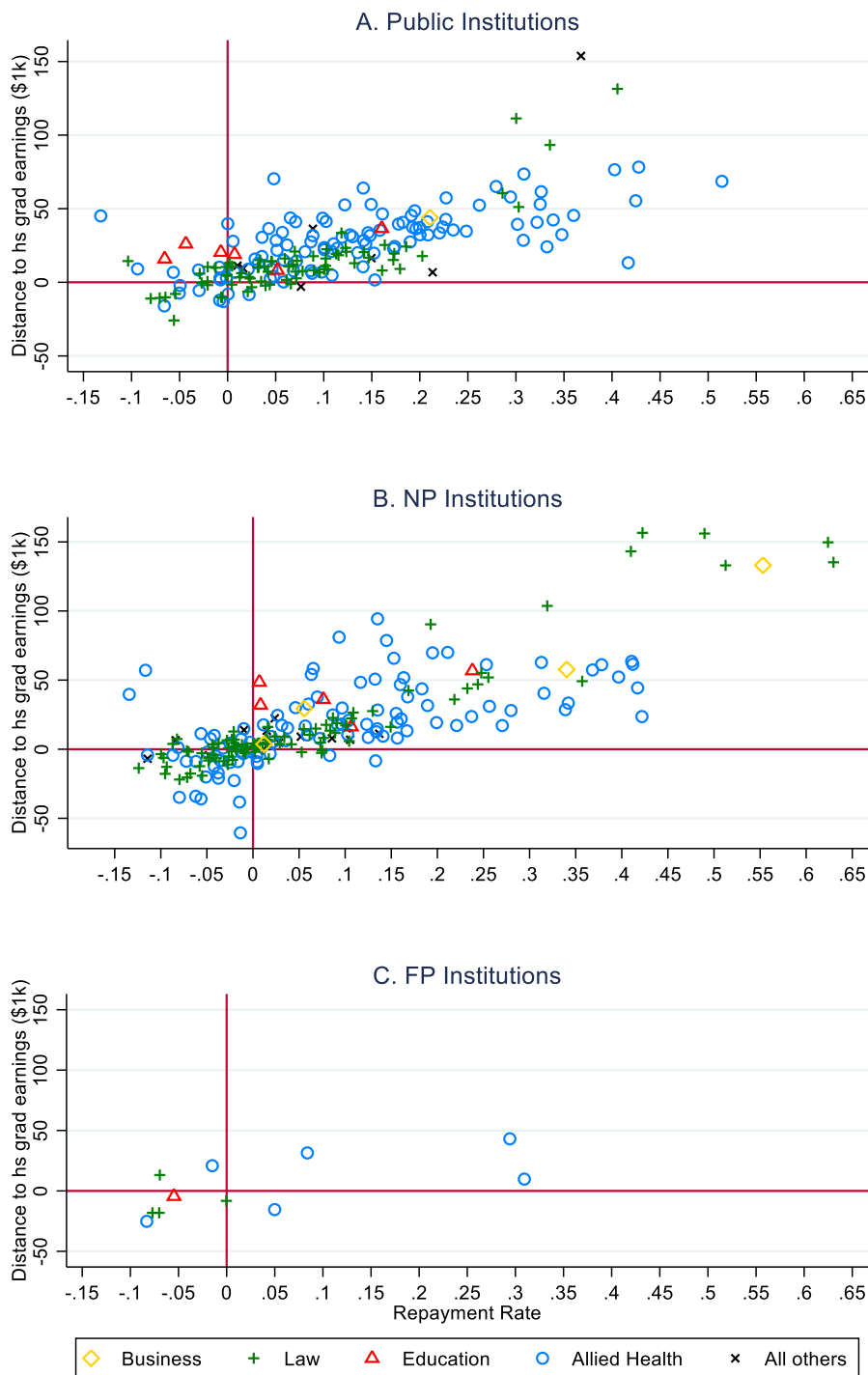|	(1) Undergrad. certificate	(2) Associate degree	(3) Bachelor's degree	(4) Grad. certificate	(5) Master's degree	(6) Doctoral degree	(7) First prof. degree	(8) All credentials
A. Reweighted to represent exiters from all programs								
All institutions	42%	51%	19%	40%	37%	30%	30%	36%
By control								
Public institutions	45%	49%	13%	33%	27%	23%	19%	34%
Nonprofit institutions	46%	60%	17%	32%	34%	28%	38%	29%
For-profit institutions	36%	65%	73%	81%	79%	60%	67%	56%
B. Unweighted								
All institutions	37%	47%	20%	37%	39%	32%	31%	32%
By control								
Public institutions	40%	47%	21%	36%	39%	31%	20%	34%
Nonprofit institutions	25%	49%	17%	32%	38%	31%	37%	24%
For-profit institutions	34%	58%	64%	74%	77%	64%	69%	47%

Notes: The loan repayment rate metric represents the fraction of the initial cohort balance that has been repaid (where a negative repayment rate indicates the cohort's balance has increased). See Sections 4 and 5 for additional details. Only programs with a nonmissing loan repayment rate (N = 46,375) are included. Panel A estimates are weighted by the number of estimated exiters from a given program (see Section 5 and Appendix B for details). Panel B is unweighted.

Table 3: Program performance on net earnings premium and loan repayment rate metrics

	(1) Negative NEP and LRR	(2) Negative NEP, positive LRR	(3) Positive NEP, negative LRR	(4) Positive NEP and LRR	(5) All programs
All programs					
% students	9%	6%	26%	59%	100%
% programs	8%	10%	15%	67%	100%
Number of students	671,750	456,179	1,952,409	4,520,715	7,601,052
Certificate programs					
% students	13%	10%	31%	46%	100%
% programs	15%	23%	15%	47%	100%
Number of students	217,580	180,212	528,405	802,627	1,728,823
Associate degree programs					
% students	11%	6%	41%	42%	100%
% programs	11%	7%	29%	54%	100%
Number of students	217,274	131,478	851,257	857,190	2,057,200
Bachelor's degree programs					
% students	1%	2%	15%	83%	100%
% programs	1%	3%	13%	83%	100%
Number of students	17,166	58,391	372,688	2,114,278	2,562,523
Gaduate certificate programs					
% students	18%	7%	17%	57%	100%
% programs	11%	13%	17%	60%	100%
Number of students	20,275	8,507	19,793	65,363	113,938
Master's degree programs					
% students	19%	7%	15%	59%	100%
% programs	15%	12%	16%	58%	100%
Number of students	175,181	69,049	144,065	551,773	940,068
Doctoral degree programs					
% students	3%	4%	27%	65%	100%
% programs	7%	4%	17%	72%	100%
Number of students	2,881	3,547	24,703	58,779	89,910
First professional degree programs					
% students	20%	5%	11%	65%	100%
% programs	16%	5%	12%	66%	100%
Number of students	21,392	4,995	11,498	70,705	108,590

Notes: NEP = net earnings premium *and* LRR = loan repayment rate (see Table 1 and 2 notes and Sections 4 and 5 for additional details). The sample includes programs that either have a nonmissing NEP and LRR or are in schools that do not participate in federal loan programs and have a nonmissing NEP (N = 22,277) are included. The first row for each credential level displays performance weighted by the number of estimated exiters from a given program (see Section 5 and Appendix B for details) and the second row is unweighted. Programs in schools that do not participate in federal loan programs (N = 445) are classified as having a positive LRR if the program's NEP is positive and as having a negative LRR if the program's NEP is negative.

Table 4: Percent, number, and costs of students in failing programs by credential level

	(1) Undergrad. certificate	(2) Associate degree	(3) Bachelor's degree	(4) Grad. certificate	(5) Master's degree	(6) Doctoral degree	(7) First prof. degree	(8) All credentials
A. Percent of enrollment in failing programs								
By sector								
Public institutions	9%	11%	0.4%	9%	11%	2%	10%	7%
Nonprofit institutions	19%	10%	1%	4%	13%	5%	27%	8%
For-profit institutions	18%	11%	2%	66%	64%	3%	65%	20%
All institutions	13%	11%	1%	18%	19%	3%	20%	9%
B. Students in failing programs and their costs and debt								
Total students	217,542	217,258	17,166	20,275	175,181	2,881	21,392	671,695
Total fed loans ($1b)								
At repayment entry	$1.472	$2.276	$0.506	$0.105	$12.355	$0.440	$2.214	$19.369
3 years after exit	$1.508	$2.351	$0.524	$0.110	$12.920	$0.458	$2.318	$20.189
Total OOP costs ($1b)	$2.389	$1.152	$0.443	$0.998	$8.306	$0.324	$3.488	$17.100

Notes: Failing programs are those that either have both a negative NEP and LRR or are in schools that do not pariticpate in federal loan programs and have a negative NEP (see Table 1 and 2 notes and Sections 4 and 5 for additional details on metrics). OOP = out-of-pocket costs.

Table 5: Distribution of institutions by the percentage of undergraduate and graduate students in failing programs

	(1) All institutions	(2) Public institutions	(3) Nonprofit institutions	(4) For-profit institutions
A. Percent of undergraduates in failing programs				
0%	0.81	0.77	0.95	0.76
1-25%	0.06	0.10	0.03	0.03
25-50%	0.03	0.05	0.01	0.02
50-75%	0.02	0.03	0.00	0.01
75-99%	0.01	0.03	0.003	0.01
100%	0.06	0.02	0.01	0.17
B. Percent of graduate students in failing programs				
0%	0.64	0.65	0.65	0.37
1-25%	0.16	0.21	0.13	0.10
25-50%	0.06	0.06	0.06	0.08
50-75%	0.03	0.03	0.03	0.06
75-99%	0.02	0.03	0.01	0.12
100%	0.09	0.03	0.12	0.29

Note: Panel A is limited to institutions that had at least 1 undergraduate program with a nopnmissing NEP and LRR or, in schools that don't participate in federal loan programs, at least 1 undergraduate program with a nonmissing NEP (N =3685). Panel B is limited t o institutions that had at least 1 graduate program with a nonmissing NEP and LRR (N = 1229). 3858 unique institutions represented across both panels. Characteristics of the remaining 1663 institutions with missing metrics are shown in Appendix Table A9.

Table 6: Distribution of Minority Serving Institutions, Historically Black Colleges and Universities, and community colleges by the percentage of undergraduate and graduate students in failing programs

	(1) MSIs	(2) HBCUs	(3) Community colleges
A. Percent of undergraduates in failing programs			
0%	0.76	0.53	0.68
1-25%	0.12	0.27	0.12
25-50%	0.05	0.14	0.08
50-75%	0.02	0.04	0.05
75-99%	0.03	0.03	0.04
100%	0.02	0	0.03
B. Percent of graduate students in failing programs			
0%	0.55	0.24	--
1-25%	0.25	0.03	--
25-50%	0.08	0.11	--
50-75%	0.03	0.11	--
75-99%	0.01	0.27	--
100%	0.08	0.24	--

Note : Limited to institutions that had at least 1 undergraduate (Panel A) or graduate (Panel B) program with a nopnmissing NEP and LRR or, in schools that don't participate in federal loan programs, at least 1 program with a nonmissing NEP. Column 1 (MSI) N = 525 (Panel A), 220 (Panel B). Column 2 (HBCU) N = 74 (Panel A), 37 (Panel B). Column 3 (CC) N = 980.

Table 7: Characteristics of institutions by the percent of undergraduate enrollment in failing programs

	% of undergrad. enrollment in failing programs				
	0%	1-25%	25-75%	75-99%	100%
12-month enrollment					
Undergraduate	5,362	22,088	7,213	7,131	815
Graduate	2,587	4,311	964	43	22
Student financial aid					
% Pell (ever)	0.67	0.81	0.86	0.88	0.86
% loans (ever)	0.79	0.65	0.57	0.51	0.81
Race/ethnicity of undergraduate student body					
% AIAN	0.01	0.01	0.01	0.02	0.01
% Asian	0.04	0.03	0.02	0.03	0.04
% Black	0.14	0.25	0.23	0.19	0.33
% Hispanic	0.16	0.16	0.19	0.19	0.17
% NHPI	0.003	0.003	0.003	0.01	0.002
% White	0.55	0.45	0.47	0.47	0.40
% URM	0.31	0.42	0.43	0.40	0.51
% Unknown or international	0.10	0.11	0.09	0.11	0.06
Student demographics					
Average age at entry	24.8	25.5	25.8	25.1	27.2
% age > 24 at entry	0.36	0.45	0.47	0.43	0.52
% female	0.63	0.61	0.60	0.64	0.69
% first generation	0.41	0.47	0.50	0.49	0.55
% veterans	0.02	0.01	0.02	0.01	0.03

Note: Limited to institutions that had at least 1 undergraduate program with a NEP and LRR or, in schools that don't participate in federal loan programs, at least 1 undergraduate program with a NEP (N =3685). AIAN = American Indian and Alaska Native. NHPI = Native Hawaiian and Pacific Islander. URM = under-represented minority (AIAN, Black, Hispanic, or NHPI).

Table 8: Institutional sector by percent of enrollment in failing programs

	A. % of undergrad. enrollment in failing programs					B. % of graduate enrollment in failing programs				
	0%	1-25%	25-75%	75-99%	100%	All pass	1-25%	25-75%	75-99%	All fail
By control and predominant degree										
Public institutions										
Certificate	0.09	0.24	0.27	0.23	0.05	0	0	0	0	0
Associate	0.13	0.30	0.43	0.49	0.06	0	0	0	0	0
Bachelor's	0.16	0.15	0.03	0.02	0	0.36	0.50	0.33	0.46	0.13
Graduate only	0	0	0	0	0	0.01	0	0.01	0	0
Nonprofit institutions										
Certificate	0.02	0.005	0.02	0	0.01	0	0	0	0	0
Associate	0.02	0.04	0.01	0.02	0.01	0	0	0	0	0
Bachelor's	0.28	0.08	0.06	0.04	0.02	0.50	0.46	0.53	0.27	0.39
Graduate only	<0.01	0	0	0	0	0.09	0.02	0.04	0	0.28
For-profit institutions										
Certificate	0.26	0.10	0.13	0.17	0.81	0	0.01	0	0	0.03
Associate	0.03	0.05	0.04	0.04	0.04	0.004	0	0	0.04	0.01
Bachelor's	0.01	0.04	0.02	0	0	0.02	0.02	0.06	0	0.04
Graduate only	0	0	0	0	0	0.01	0	0	0	0.06
HBCUs	0.01	0.05	0.16	0.05	0.01	0.01	0.05	0.16	0.05	0.01
MSIs	0.10	0.22	0.19	0.42	0.04	0.10	0.22	0.19	0.42	0.04

Note: Panel A (B) is limited to institutions that had at least 1 undergraduate (graduate) program with a NEP and LRR or, in schools that don't participate in federal loan programs, at least 1 undergraduate (graduate) program with a NEP (Panel A N = 3685 institutions, Panel B N = 1229 institutions, 3858 unique institutions represented in both panels). HBCU = Historically Black College or University. MSI = Minority Serving Institution.

Table 9: Undergraduate outcomes, institutional finances, and institutional accountability performance by percent of undergraduate students in failing programs

	% of undergrad. enrollment in failing programs				
	0%	1-25%	25-75%	75-99%	100%
Undergraduate outcomes					
% with earnings > $25k @ 10 years	0.68	0.59	0.55	0.55	0.36
% paying down $1 student loan debt @ 3 years	0.53	0.35	0.33	0.32	0.28
% paying down $1 student loan debt @ 7 years	0.62	0.47	0.43	0.44	0.36
Cohort default rate @ 3 years	0.09	0.14	0.16	0.16	0.15
Income-driven repayment participation					
% borrowers	0.20	0.22	0.21	0.20	0.20
% balance	0.28	0.32	0.30	0.28	0.26
Institutional finances					
Tuition/FTE	$11,303	$6,544	$5,168	$5,009	$8,338
Instructional expenditures/FTE	$8,835	$6,656	$5,815	$5,848	$4,509
Institutional accountability					
Loan repayment rate	0.08	0.002	-0.01	-0.01	-0.03
Negative loan repayment rate	0.13	0.51	0.68	0.71	0.95
Net earnings premium	$11,275	$7,094	$3,758	-$354	-$5,232
Negative net earnings premium	0.23	0.20	0.52	0.86	1.00

Note: Limited to institutions that had at least 1 undergraduate program with a NEP and LRR or, in schools that don't participate in federal loan programs, at least 1 undergraduate program with a NEP (N =3685). Institutional repayment rate is defined similarly to program-level loan repayment rate. Institutional net earnings premium is the exiter-weighted average of program-level net earnings premia.

Table 10: Aggregate costs of no accountability for students in failing programs with passing alternatives in the same institution

	All institutions		Public institutions		Nonprofit institutions		For-profit institutions	
	(1) Failing programs	(2) Alternatives	(3) Failing programs	(4) Alternatives	(5) Failing programs	(6) Alternatives	(7) Failing programs	(8) Alternatives
A. Undergraduate students								
Earnings	$17,357	$30,161	$17,118	$29,782	$17,547	$29,416	$18,099	$31,837
OOP costs	$8,788	$10,039	$4,840	$5,083	$22,035	$31,561	$15,440	$16,745
Net earnings premium	-$2,707	$10,010	-$2,804	$9,843	-$2,935	$8,316	-$2,261	$11,392
Negative net earnings premium	1	0.05	1	0.06	1	0.01	1	0.04
Median student loan debt	$16,021	$18,127	$13,968	$14,352	$19,901	$27,697	$21,350	$26,780
Average repayment rate	-0.04	0.003	-0.04	0.01	-0.04	0.001	-0.04	-0.02
Negative repayment rate	1	0.57	1	0.50	1	0.65	1	0.77
Average # of programs in school								
Failing programs	1.8		1.9		1.5		1.4	
Passing programs	9.8		8.8		12.7		11.9	
Programs with missing metrics	19.6		24.1		12.1		7.2	
Total failing program exiters	370,643		271,478		47,939		51,225	
Tot. loss in net earnings ($1m) per cohort	-$4,713		-$3,433		-$539		-$699	
Total debt accumulated in failing progs ($1m)	$5,938		$3,792		$954		$1,094	
B. Graduate students								
Earnings	$38,438	$59,582	$35,388	$59,662	$38,605	$63,628	$43,797	$53,871
OOP costs	$56,349	$52,567	$45,947	$46,118	$76,531	$59,340	$47,646	$55,065
Net earnings premium	-$6,624	$12,719	-$6,355	$13,777	-$7,838	$15,409	-$5,448	$7,077
Negative net earnings premium	1	0.08	1	0.12	1	0.08	1	0.01
Median student loan debt	$56,349	$52,567	$45,947	$46,118	$76,531	$59,340	$47,646	$55,065
Average repayment rate	-0.04	0.06	-0.04	0.08	-0.04	0.07	-0.04	-0.01
Negative repayment rate	1	0.32	1	0.20	1	0.32	1	0.53
Average # of programs in school								
Failing programs	2.9		2.1		2.1		5.3	
Passing programs	7.2		7.8		7.5		5.9	
Programs with missing metrics	22.6		32.5		21.3		6.3	
Total failing program exiters	190,615		55,590		49,454		85,571	
Tot. loss in net earnings ($1m) per cohort	-$3,687		-$1,119		-$1,150		-$1,072	
Tot. increase in debt ($1m)	$10,741		$2,554		$3,785		$4,077	

B | Economic Studies
at BROOKINGS

The Brookings Economic Studies program analyzes current and emerging economic issues facing the United States and the world, focusing on ideas to achieve broad-based economic growth, a strong labor market, sound fiscal and monetary policy, and economic opportunity and social mobility. The research aims to increase understanding of how the economy works and what can be done to make it work better.