Reforms to Increase Transparency in Higher Education

Mark S. Schneider
Vice President, American Institutes for Research

Higher education is one of the biggest investments that individuals make over the course of a lifetime. To help students make the most of this investment, federal higher-education policy supports portable grants, loans, and tax credits available to prospective students and allows them to choose from a diverse array of providers. When the system was designed, policymakers assumed that providing voucher-like Pell grants, for example, and future tax benefits to students and allowing them to choose would reward schools that offer high-quality programs and punish those that fall short. In the aggregate, it was hoped, these choices would create market forces that would hold colleges and universities accountable for what they charge and the quality of the education they deliver.

Market competition works best when consumers can find and use clear, comparable information about the costs and quality of different offerings. If such information is lacking, either because it does not exist or because it is difficult to find and use, then market competition will be based on other attributes that may or may not be related to the key dimensions that enhance quality and efficiency. In the case of higher education, that means students might judge campuses based on their proximity to home, amenities (lazy rivers, climbing walls, top chefs), or, in some cases, tuition costs (as a proxy for quality). In the aggregate, choices based on these dimensions might reward campuses that have a geographic monopoly or those that inflate their tuition, stunting the ability of market forces to improve the system as a whole.

To be sure, evaluating the quality of post-secondary institutions and programs is a difficult task, even when information is plentiful. Part of
this is because of the nature of the good: A post-secondary education is an “experience good,” meaning it is difficult to assess a school’s value until after you’ve actually enrolled. In some cases, the true value is not recognized until many years in the future when graduates learn how much their degree is rewarded in the labor market. And most students only purchase a post-secondary education once or twice, meaning they have little opportunity to learn from experience.

Consumers also face a dearth of clear, comparable data on the cost and quality of different offerings. Some basic pieces of information, such as the actual out-of-pocket costs for a given student at a given institution, are available only at the very end of the college-application process, after students have settled on a set of choices (and schools often change the terms of their financial-aid packages from year to year).

Other information is incomplete: Federal graduation rates, which provide a basic measure of the likelihood of completing a credential, are currently based on first-time, full-time students only, which excludes students who transfer in and complete a credential or transfer out and complete one somewhere else.¹⁸⁸ Data on how much students learn is largely non-existent. And information on how graduates of particular programs fare after finishing school—in terms of finding a fulfilling job, repaying loans, and contributing to society—is also not systematically available outside of a handful of states or institutions. Popular private rankings suffer from the same limitations.

The federal government, in concert with the states and institutions, could do more to increase transparency and enhance market accountability in higher education. More effectively reporting data that it already collects and collecting better data on basic measures of cost, quality, and value would provide a number of benefits.

First, students could use the information to avoid investing in schools or programs that do not provide a positive return on investment and to discover options that they may have eliminated on the basis of incomplete or faulty information. For instance, while many argue that a bachelor’s degree is the only surefire path to the middle class, a closer look at the earnings of workers with associate’s degrees or certificates in technical fields, or those who complete apprenticeships, reveals that there are many other affordable, worthwhile opportunities to consider.¹⁸⁹

Second, researchers and policymakers could more readily judge where investments in federal aid are paying off and where reforms
could improve efficiency and reduce waste. Though the Office of Federal Student Aid sits on millions of student-level records that measure the receipt of grants and loans, completion or separation status, and loan repayment, very little of that data is used to inform the policymaking or budgeting process. And almost none of those administrative data are made available to researchers who could help answer pressing questions.

Third, private firms could use new, more granular data to come up with all manner of rankings and ratings to reflect the unique preferences of different students. The most popular rankings tend to reward admissions selectivity and spending over actual measures of student learning or value-added. Better data on post-graduation outcomes would provide a fuller picture of institutional quality and, eventually, encourage institutions to compete on how well their graduates do after graduation rather than how well they scored on their entrance exams. Early evidence suggests that the earnings data released on the newly revamped College Scorecard affected student choices.¹⁹⁰

Fourth, private lenders and funders could use labor-market outcome data to improve underwriting and extend credit on the basis of a student’s potential rather than the student’s past experience with credit products. Without reliable data on the likely return on investment to different options, lenders are forced to rely on credit scores and the availability of credit-worthy co-signers. These measures exclude students who may have high potential but no credit history.¹⁹¹

While there are opportunities to enhance transparency, it is important to place clear restrictions on what federal regulators can use such data for, to make sure these efforts are designed to serve a specific audience and to protect students’ privacy.

**The Status Quo in Federal Data Collection**

While slow and “under the radar” changes are always occurring in federal data collection, perhaps the most visible attempt to rewrite the federal role was the Obama administration’s failed attempt to build a Postsecondary Institutional Rating System (PIRS). This effort serves as an important reference point for assessing the challenges and opportunities facing federal data efforts in post-secondary education.

In 2013, the White House decided that the nation needed a rating system that would evaluate the approximately 7,000 post-secondary
Policy Reforms to Strengthen Higher Education

institutions that participate in federal student-aid programs. To its detriment, PIRS was straddling two different tasks from the very beginning. On the one hand, PIRS was going to produce data and a data tool that could be used by consumers to evaluate the quality, accessibility, and affordability of post-secondary institutions, allowing them to make better-informed choices. But, as a second goal, and far too ambitiously, the White House sought to link performance on PIRS with eligibility for continued Title IV funding.

The tie to Title IV funding was inevitably abandoned and the College Scorecard, when released, was an adequate, but not great, consumer information site. Perhaps the most important aspect of the rollout of the Scorecard was the fact that the Department of Education made public a large database, only some of which was actually used in the Scorecard itself.¹⁹²

The explicit purpose of this data release (with accompanying instructions on how to access it using APIs¹⁹³) was to allow researchers and others easy access to a treasure trove of data. This was an acknowledgement that the Department of Education, while having perhaps a unique ability to collect data, did not have a unique ability to deploy it.

There are at least three lessons from this effort that should be kept in mind as any new administration approaches the need for better data on post-secondary institutions. First, the power of existing institutions of higher education — and their top lobbying organization, the American Council on Education (ACE) — is formidable. While not as powerful or uniformly opposed to good data and good measurement as in the past, the opposition of the higher-education “industry” to ratings and to the administration’s plan to tie ratings to Title IV funding helped weaken PIRS.

Second, the federal government must be careful about mixing consumer information tools and regulatory tools. While there may be overlap in the information consumers need and the information policymakers need, mixing the two creates problems. And the way in which data are collected, curated, and displayed varies greatly depending on the primary focus of the effort.

Finally, we must recognize that the data that the federal government has to measure student outcomes is limited. Ultimately, the success of students and institutions should be measured by how much students earn after they leave school and how much they learned while
attending. There is some agreement on assessing labor-market outcomes. In contrast, there is little to no agreement on how to measure what many would call the most basic product of higher education: student learning. For instance, a recent report by ETS argued that there is a need for a “systematic, data-driven, comprehensive approach to understanding the quality of post-secondary education...with direct, valid, and reliable measures of student learning.” In that report, ETS explores the challenges of creating such a measurement system—including the difficulty of defining the different dimensions that should be included in such a measure of student learning, ranging from workplace skills to academic expertise and encompassing both “hard skills” as well as so-called “soft skills,” such as teamwork and creativity.¹⁹⁴ Given the breadth of these different demands, little consensus now exists on how to move forward. In turn, it is probably misguided for the federal government to invest scarce time and resources at this point trying to develop measures of learning outcomes for post-secondary education.¹⁹⁵

The absence of data on student learning and the relative paucity of good data on student earnings highlight the limits of federal data systems. At the same time, an established infrastructure for other measures of student success exists with clearer means for improvement.

To start, the federal government’s primary source of data on post-secondary education is the Integrated Postsecondary Education Data System (IPEDS), which requires institutions that participate in federal student-aid programs to fill out a series of surveys each year. The surveys focus on 12 distinct topics, including the following: institutional characteristics, institutional prices, admissions, enrollment, student financial aid, degrees and certificates conferred, student persistence and success, and institutional resources.¹⁹⁶ This extensive coverage of so many aspects of higher education—the topics covered, the very questions asked, and the mixing of consumer and regulatory information—is the result of a long process of accretion whereby legislation demands that new pieces of data be collected but never acts to eliminate questions or whole surveys that have outlived their usefulness (if they ever had any to begin with).

In IPEDS, the collected data are aggregated at the institution level, providing a snapshot of an institution’s enrollments, finances, staffing, prices, and some student outcomes in a particular year. IPEDS is the only source of comparable institution-level data on student outcomes like retention and graduation rates. Much of IPEDS data are extensive but flawed. For
example, graduation-rate data have historically been based on the cohort of first-time, full-time beginning students, therefore lacking any ability to track the success of students who transfer. The finance-data collection is also of limited utility, since the data are at the institution level and not related to activities or costs associated with those activities. The Human Resources survey is hardly ever used — and almost universally nominated as the most expendable of the IPEDS surveys.

In addition to formal reporting requirements, institutions must disclose information on a number of topics to prospective students and the public. The latest reauthorization of the Higher Education Act (in 2008) contained 40 separate disclosures (nine of which only had to be disclosed to loan borrowers).¹⁹⁷ However, there is evidence that compliance with those disclosure requirements is spotty.¹⁹⁸

Disclosure requirements range from essential aspects of institutional activity — student financial-aid information, student outcomes, and health and safety — to peripheral aspects — availability of voter-registration forms and information about intercollegiate athletic programs. The disclosure requirements are often extensive and detailed. Take, for example, the disclosure requirements regarding student financial-aid information:

Each institution must notice all enrolled students of all the need-based and non-need-based federal, state, local, private, and institutional student financial assistance programs available to students who enroll in the institution; the terms and conditions of all available federal loans; the criteria for selecting recipients and for determining amount of award; eligibility requirements and procedures for applying for aid; methods and frequency of disbursements of aid; rights and responsibilities of students receiving federal student aid, including criteria for continued student eligibility and standards for satisfactory academic progress; terms of any loan received as part of the financial aid package, sample loan repayment schedule, and the necessity for repaying loans; a statement that enrollment in a program of study abroad approved for credit by the home institution may be considered enrollment in the home institution for purposes of applying for federal student financial aid; general conditions and terms applicable to employment provided as part of financial aid package; the exit counseling information the institution provides and collects.¹⁹⁹
In addition to institution-focused data efforts, the National Center for Education Statistics (NCES) conducts a handful of surveys of representative samples of students. These surveys are a key source of information about post-secondary education trends and the determinants of college access and success. Given the large federal investment in student aid, the National Postsecondary Student Aid Survey (NPSAS), which examines how students and their families pay for post-secondary education, is likely the most important.

NPSAS includes nationally representative samples of undergraduate, graduate, and first-professional students in public and private institutions ranging from community colleges to major research universities. Students who do not receive financial aid also are included in NPSAS. NPSAS is usually conducted on a four-year cycle and combines student surveys with administrative records. It is a good overall survey and can be disaggregated to different sectors (e.g., public or private) and different levels (e.g., community colleges or four-year colleges). But NPSAS cannot provide any information on individual institutions or even smaller categories of schools (e.g., Ivy League, community colleges in New York). And it is administered as one giant survey usually every four years. This time gap can be too long to capture fast-changing economic conditions.²⁰⁰ Recently, NCES announced a new, biennial NPSAS Administrative Collection, which will allow for representative financial-aid estimates on the national and state levels. The data will be made available to researchers in 2019.²⁰¹

NPSAS provides the sampling frame for longitudinal studies, such as the Beginning Postsecondary Students Longitudinal Study (BPS). BPS surveys cohorts of first-time, beginning students at the end of their first year and then three and six years after starting post-secondary education. It collects data on items such as student demographics, school and work experiences, persistence, transfer, and degree attainment.²⁰²

NPSAS also provides the sampling frame for the Baccalaureate and Beyond Longitudinal Study (B&B). This study focuses on the education and work experiences of graduates after they complete a bachelor’s degree, with a special emphasis on the experiences of new elementary and secondary teachers. The most recent B&B cohort was drawn from the 2008 NPSAS sample and approximately 19,000 students were contacted again in 2009 and again in 2012.²⁰³

These longitudinal data are primarily designed for academic
researchers, and there is little evidence that the information contained in them is used by policymakers or by consumers.

Beyond data collection and reporting, the Department of Education also provides a number of consumer-facing tools, which have proliferated in recent years. The department has grouped many of them under one umbrella: The College Affordability and Transparency Center (CATC).²⁰⁴

Links take the user to several different consumer-oriented department sites: College Navigator, College Scorecard, Net Price Calculator Center, the College Affordability and Transparency List, 90/10 Information, and State Spending Charts. Many of these linked sites sit on top of more complicated databases that can help (savvy and motivated) consumers investigate various aspects of colleges with varying degrees of ease. The College Navigator is a semi-user friendly interface to IPEDS data; the State Spending Charts display a very limited slice of IPEDS expenditure/revenue data, but the intended audience seems a bit opaque. The 90/10 Information center lists all proprietary schools that receive more than 90% of their funds from the federal government—though it is entirely unclear how a prospective student would use such information in making a choice.

Consumers actually use some of these resources, especially the College Navigator and the College Scorecard. But these tools are no better than their underlying data. As noted above, much of the IPEDS data is fundamentally flawed, and the College Scorecard earnings data are arguably at the wrong level of aggregation: They report earnings at the institution level, when there is far more variation across different departments within a college than across colleges.²⁰⁵

The Education Department also maintains federal student-aid databases. As a function of administering the student-loan programs, the department’s Office of Federal Student Aid (FSA) tracks student-loan borrowers via the National Student Loan Data System (NSLDS), a student-level database. The database provides information to students on the status of their federal aid. FSA also has a data center that provides public information on the federal student-loan program, including the aggregate performance of the federal student-loan portfolio, institutions’ loan performance (loan volume per institution, cohort default rates, gainful-employment information), and reports on lenders and guaranty agencies. FSA also uses the NSLDS to monitor institutional compliance with federal aid programs.²⁰⁶
The NSLDS contains five main types of data: demographic information about a student, records of a student’s financial-aid history (types of aid, when they received it, and college attended), information on when a student left a school (graduated or withdrew), information about student-loan repayment (loan servicer, loan status, and outstanding balance), and information from the FAFSA and the student’s dependency status.²⁰⁷

There are two other FSA databases: the Central Processing System (CPS) and the Common Origination and Disbursement (COD) system. CPS stores information about student-aid applicants such as dependency status, parental information, and income, as well as the calculations run on a student’s aid application—namely the expected family contribution (EFC). The COD system assists in sending aid money to schools; it stores data on disbursement amounts and student demographics. Data in CPS and COD link with NSLDS; the latter serves as long-term storage for all information on federal financial aid.²⁰⁸

Because the FSA data are at the student level, they potentially have far more value than the data collected by IPEDS. But there is a fundamental difference in how the data are handled by FSA and NCES. As a federal statistical agency, NCES has in its mission and culture the goal to share its data as widely as possible, while still protecting the privacy of our nation’s citizens. FSA has been classified as a Performance Based Organization (PBO), focused on administering aid programs, not reporting data or facilitating research. In other words, FSA is essentially a bank, and its culture does not support expanding access to its data.

Areas for Improvement

Despite being the object of constant tinkering, federal data collection and reporting still fail to answer basic questions regarding the purpose of higher education—and regarding the return on the huge investments made by students and taxpayers. Among the most important questions that are still difficult to answer: What can students expect to pay out of pocket (the net price after grants and scholarships)? How do students fare in the labor market after leaving college, and what’s their return on investment? How do they fare on other important metrics of student success (student persistence, degree completion, loan repayment)?

The first challenge is measuring the price of an education. Though the sticker price of tuition tends to garner the most media coverage, only a
fraction of students actually pay that price. Most pay less thanks to grants, scholarships, and tuition discounts. Institutions use their access to fine-grained financial information to price discriminate, or tailor tuition prices to individual students’ ability and willingness to pay. As such, sticker prices (which are readily available) might be misleading vis-à-vis what any given student will actually be charged. Unfortunately, most students don’t know what they’ll actually have to pay until after they’ve applied for admission and financial aid, been accepted, and received a financial-aid offer letter. In other words, they don’t know their real out-of-pocket costs until long after they’ve narrowed their choices to a handful of institutions.

The federal government has made progress on this front, requiring each college that receives Title IV aid to create a “net price calculator” that provides students with an estimate of what students who share their income and academic profile actually pay. Institutions are also required to report average net prices by income quartile as part of the IPEDS survey. However, both of these efforts are flawed. The net-price estimates in IPEDS are based on students who received any grant aid, which leaves out large numbers of students in the middle and higher income groups.²⁰⁹ They are also averages, meaning fluctuations from year to year may reflect changes in the distribution of students within an income category rather than changes in the net price any one student is charged.²¹⁰ And early analyses of net-price calculators have found significant variation in the information required from students and in the way net prices are displayed, making it difficult for students to compare across colleges.²¹¹

Another troubling fact is that very few institutions make multi-year commitments of financial aid (or of a fixed tuition cost). Sometimes financial aid is “front loaded”—offering an attractive first-year award to lure a student to register.²¹² Even if a student is making good academic progress, increases in tuition or decreases in financial aid can change the out-of-pocket (net) costs of attending—and usually not in a direction favorable to the student. Information on this is close to non-existent.

The next challenge is measuring the earnings of students after they leave school. While the federal government collects lots of data on post-secondary education—and even though the College Scorecard published data about the earnings of students enrolled in post-secondary institutions six and 10 years after enrolling, much of the data that are available to measure the labor-market success of students is inadequate. Most notably, the earnings measures in the Scorecard were based on students
who received federal financial aid, and they were also aggregated at the institutional rather than the program level. As a result, the main source of post-secondary earnings data does not measure much of the variation in outcomes. In addition, the Scorecard data lump all incoming Title IV students together, which does not distinguish between students who completed credentials and those who did not.

As a result, we know very little about how students from different institutions and different programs of study fare after college. This makes it impossible to adequately measure the return on investment (ROI) of students or taxpayers, raising significant questions about what we are actually getting for the billions of dollars that the federal government, state governments, and families invest in post-secondary education. While we know that, on average, post-secondary education is a good investment, ROI varies widely across colleges and universities — and even more across different fields of study.²¹³

To measure ROI at the institution and program level, one would need to merge two different sets of data. The first are individual student-level “transcript” data that show the year a student completed a course of study, the institution that awarded the post-secondary credential, and the field of study (this is the federal Classification of Instructional Program code, known as the CIP code). The second are wage data. At present, these wage data mostly come from state unemployment insurance (UI) wage systems, although the Scorecard used the more comprehensive unduplicated W-2 wage data from the IRS.

Merging student-level data with either source of wage data uses Social Security Numbers, and the merging is usually done by the agency that holds the wage data (to protect privacy). The individual-level data are never made public. Rather the data are aggregated at the program level, inspected to suppress any small programs (as a rule of thumb, programs that contain fewer than 10 cases are suppressed), and returned to the education agency that provides the transcript-level data.

There are currently no nationwide standards governing how these data are used. For example, to minimize the number of missing programs caused by small enrollments, states that release merged transcript/wage data often combine several cohorts. Practices across states differ somewhat, but this is a technical issue that could (and should) be resolved by the federal government.

There is also a question about what to do with students who enroll
Policy Reforms to Strengthen Higher Education

in but do not complete a program. Most states are focused on the wages of completers, but, as is well known, large numbers of students never finish. The federal Scorecard data tracked cohorts of students, but did not distinguish between completers and leavers. The transcript data can also include demographic information (e.g., race or gender). This could provide valuable information about the differential success of different types of students, but adds complexity to the aggregated data.

Yet another challenge is the level of data needed by the federal government to assess student success. As noted, the Scorecard used data only on students who participated in a Title IV program. Because the Department of Education must know whether or not students are in good standing with an institution of higher education in order to know when students must begin repaying their loans, the NSLDS maintains detailed records of the enrollment of students receiving federal aid in any Title IV approved institution. This effectively creates a student-level data system for the majority of students in the nation—despite the existing ban on the federal government holding such data. Moreover, Title IV student-level data actually chart the path of the students in which the nation’s taxpayers are investing the most money. And there is certainly a compelling federal interest in knowing the extent to which Title IV students are succeeding in the pursuit of post-secondary credentials.

The federal Scorecard only reported wage data at the institution level, the only level at which the NSLDS can currently collect data. The Department of Education may overcome this flaw in the next several years because institutions must now report to FSA information on the programs in which students are enrolled. (This information is needed because the 150% Subsidized Loan Limitation provisions are based on the borrower’s enrollment in a specific program.) Because student outcomes vary greatly across programs of study both within and across institutions, these program-level data are essential. In short, to the extent to which FSA collects student-level indicators of success at the program level for students who have received federal student loans and/or Pell Grants, the nation has the potential to better measure the payoff of the large investment the nation is making in its post-secondary students.

While these data improvements are taking hold at FSA, the venerable and outdated IPEDS data system is also expanding its measurement of student outcomes. Starting in the 2015-2016 academic year, institutions will have to report more detailed information about the success of transfer and
part-time students. Schools will be required to report completion data for four cohorts of degree/certificate-seeking students: full-time, first-time students; part-time, first-time students; full-time, non-first-time entering students; and part-time, non-first-time entering students. These data are valuable and IPEDS should continue to refine and expand its student-outcome measures. But the changes in FSA data, if properly shared with the public, can dwarf the benefits of the expanded IPEDS data.

As noted above, FSA has been classified as a PBO since the 1998 reauthorization of the Higher Education Act. Its orientation is essentially that of a bank, focused solely on the administration of financial aid programs rather than reporting data or facilitating research. Title I, Part D of the 1998 HEA lays out seven priorities for FSA as a PBO (hereafter called “Purposes as a PBO”):

A. to improve service to students and other participants in the student financial assistance programs authorized under subchapter IV of this chapter and part C of subchapter I of chapter 34 of title 42, including making those programs more understandable to students and their parents

B. to reduce the costs of administering those programs

C. to increase the accountability of the officials responsible for administering the operational aspects of these programs

D. to provide greater flexibility in the management and administration of the Federal student financial assistance programs

E. to integrate the information systems supporting the Federal student financial assistance programs

F. to implement an open, common, integrated system for the delivery of student financial assistance…

G. to develop and maintain a student financial assistance system that contains complete, accurate, and timely data to ensure program integrity.

Under its current mandate, FSA is primarily concerned with its core jobs: assessing eligibility for aid, disbursing the aid, and tracking repayment. FSA is required to report some basic data on loan-default rates, and its data center provides access to aggregate data on loan
Policy Reforms to Strengthen Higher Education

disbursements; the distribution of repayment plans; the frequency of forbearance, deferment, and delinquency; and institution-level data on defaults, program reviews, and financial responsibility scores.²¹⁶ However, as a PBO, FSA has often been less than responsive to requests for data and research that would benefit the rest of the nation. This presents a clear opportunity for reform.

**Improving the Current System**

There are several paths potential reformers could take to improve transparency in our higher-education system. The most modest of these paths should start with the FSA, by inserting new goals into FSA’s “Purposes of the PBO” that call for a more active role in reporting on NSLDS data, assessing the effectiveness of federal investments, and facilitating research.

Higher-education observers have long argued that FSA should be more engaged in data reporting and research. Inserting such goals can help reformers enlist FSA and its wealth of data in the effort to boost transparency. While its role as a bank and originator of direct federal student loans must remain paramount, its structure as a PBO provides an opportunity to make FSA more responsive to the dissemination of data. Specifically, the chief operating officer must create an annual performance plan for FSA in consultation with students, institutions, Congress, lenders, and others. That plan should include the development and dissemination of data measuring the results of the taxpayers’ $130 billion annual investment in student financial aid. A formal revision of FSA’s “Purposes as a PBO” could make this a core part of FSA’s mission.

For starters, point (G) could be revised to include other uses for FSA data besides program integrity, such as “to develop and maintain a student financial assistance system that contains complete, accurate, and timely data to provide updates on the state of the federal loan portfolio, assess the effectiveness of federal investments, and ensure program integrity.” Reformers could also add a requirement that FSA take an active role in informing the policymaking process in Congress and the executive branch.²¹⁷ Finally, reformers might consider adding a goal related to producing or facilitating new research using FSA data.²¹⁸ In exchange for these additions, reformers could delete goals (E) and (F), which seem to be outdated.

Next, reformers should revise the College Scorecard to fix its flaws and improve its accuracy, and then commit to maintaining it. The
Obama Administration created the College Scorecard in 2013 and revamped it dramatically in 2015. By then, the administration had abandoned its proposal to create a system of federal college ratings and instead released a host of previously unavailable outcomes data on the Scorecard. The Department of Education released institution-level data on loan-repayment rates, alumni earnings at various points up to 10 years after enrollment, and the proportion of alumni earning less than $25,000 a year. To produce the earnings data—the first federal effort of its kind—the department worked with the IRS to match financial-aid records with wage records for particular cohorts of students.

While a positive step in many regards, the Scorecard suffers from some significant shortcomings. First, it only covers students who received federal financial aid, leaving out a substantial number of students. Second, its earnings variables aggregate those who completed a degree with those who did not, which obviously produces a misleading picture of the value of completing a degree. Third, the data are not disaggregated at the level of program of study, even though earnings vary dramatically across majors (often more than they vary across institutions). Last, because the Scorecard was an Obama administration creation (and not mandated in statute or regulation), there is no guarantee that it will be maintained or updated.²¹⁹

An incoming administration should commit to fixing the existing flaws that it can, most obviously disaggregating earnings data according to whether the student graduated from the school and according to the program of study. In addition, shifting the Scorecard to NCES would give it a permanent home with an agency that exists to provide educational data. Doing so would also enable NCES to undertake a series of technical review panels to decide, with input from the higher-education sector, how to consistently measure the metrics of interest.

Reformers should also streamline the department’s consumer-facing information resources. The College Scorecard is just one of many consumer-facing college tools that the federal government now runs. The creation and maintenance of NCES’ College Navigator is required under the Higher Education Act (20 U.S.C. 1015a), and the 2008 reauthorization of the law also required the creation of the College Affordability and Transparency Center (CATC).²²⁰ This has unfortunately added unnecessary complexity to the system, especially when different data sources may have different answers to the same question due to differences in methodology or timeframe.
The jumble of consumer-facing information sources should be simplified into one comprehensive tool with consistently defined measures on the quantities that matter. The Strengthening Transparency in Higher Education Act, passed by a bipartisan majority of the House Committee on Education and Workforce in 2014, would revise Section 132 of the Higher Education Act (20 U.S.C. 1015a) to create a single set of institutional dashboards containing relevant consumer-facing data (called the “College Dashboard”). It would also require the secretary to publish on these dashboards several additional data points: completion rates for students who receive a Pell Grant, those who take on a federal student loan, those who receive Defense Department or veterans education benefits, and those who receive no federal aid; the average federal student-loan debt of graduates who borrowed; and a link to national and regional data from the Bureau of Labor Statistics (BLS) on starting salaries in “major occupations.”

The proposed College Dashboard would take the place of the College Navigator and the CATC. The College Scorecard, as noted, is not required by statute or regulation, is not mentioned in the bill, and could remain active as long as desired by the sitting administration and Congress.

Though House Republicans are reticent to embrace the College Scorecard given its association with the Obama administration, rather than create an entirely new tool (at additional expense), the most cost-effective approach may be to build the proposed dashboard on the existing Scorecard interface (and to include some of the useful data currently on the Scorecard, like loan-repayment rates). The legislation’s other recommendations—that students be provided with a link to the dashboard of any college they name on the FAFSA—would still be possible.

As this effort goes forward, federal policymakers need to confront a basic decision: Should this be a regulator-facing dashboard or a consumer-facing one? While much of the underlying data may be the same, how the data are displayed, the ease with which users can compare (and rank) programs and colleges, and the very choice of which data to highlight in the dashboard will in part be determined by that decision. The mishmash of purposes makes for the mishmash of programs collected in the CATC.

**Further Steps**

For reformers wishing to go further, there are a series of bolder policies to pursue. Allowing researchers access to FSA data extracts would be a good place to start. The National Center for Education Statistics
(NCES) allows access to restricted-use data to qualified researchers via an established application process. Researchers can access student-level data from NCES sample surveys, which facilitates research that in turn informs our understanding of federal investments and student outcomes. Access is tightly controlled; researchers must have an approved security plan and violations of federal privacy rules are punishable by a $250,000 fine.²²² The public can access some of NCES’s data through its Datalab. Depending on the question, the public can access information using QuickStats, PowerStats, or TrendStats. Each of the three “Stats” programs (and their underlying data) have been carefully constructed so that no personally identifiable information can be uncovered — yet analysts can address valuable and important questions.

There are no such routes to de-identified data from NSLDS. Yet the few studies that have been able to use NSLDS data have been invaluable in uncovering important trends and problems in federal aid programs. One such study, which merged a sample of NSLDS data with tax-return information from the Treasury Department, found that subsets of students who attended open-access institutions were struggling mightily to pay back their federal loans, often because they failed to complete a credential. Five-year loan-default rates were far higher than published three-year rates among for-profit and community colleges.²²³ Such studies can help to inform the policymaking process by identifying problems more clearly and helping to target solutions.

In April 2016, the Department of Education started exploring ways to allow researchers access to federal data for studies that “can inform and advance policies and practices that support students’ post-secondary success and strengthen repayment outcomes for borrowers.”²²⁴ The first researchers to get access to these data will be from the Federal Reserve Board, who will be able to match student-aid data files with other data. This policy focus will hopefully leverage data to increase efficiency, transparency, and accountability.²²⁵

An incoming administration could continue these efforts by adopting some of the recommendations offered by Matthew Soldner and Colleen Campbell in a recent paper on the subject.²²⁶ First, a new administration could direct FSA and NCES to collaborate on adding de-identified data extracts from FSA’s different data systems to the list of datasets available to analysts with a restricted-use data license. Analysts could then access those data extracts via the standard application procedure. Second, a
new administration could leverage the Education Department’s newly created Enterprise Data Warehouse and Analytics (EDWA) project to facilitate researcher access. Soldner and Campbell recommend adopting the Census Bureau’s “Research Data Centers” model, where researchers can access restricted Census data only after having a research plan approved and traveling to one of dozens of secure physical sites.

Further, policymakers should allow the use of IRS Form 1098T to calculate key measures of cost and student outcomes. Form 1098T is an under-used tool for improving data and transparency, as it facilitates the claiming of higher education tax benefits. Colleges fill out a 1098T for every student they enroll that has a “reportable transaction;” the form provides the IRS with information on the student’s institution, their enrollment intensity (more than half time) and level (graduate or undergraduate), the amount of tuition paid, and the amount of scholarship aid received. Students and/or their families also receive a copy for use in filing their taxes.

Observers have noted that the 1098T could be useful in providing information on net price and, with some augmentation, program-level degree completion and program-level post-college earnings. The Department of Education could use the information reported on tuition and scholarships to calculate an average net price by income group for students who received federal aid, thus obviating the need for institutions to create their own net-price calculators.

Measuring student outcomes would require a more sizable revision. Grover “Russ” Whitehurst, founding director of the Institute for Education Sciences, has proposed the addition of two boxes to the 1098T—one that captures degree completion and one that captures the program of study (using the CIP code). Institutions already have to report both pieces of information to the Department of Education as part of Title IV compliance, so reporting them on form 1098T should not constitute an additional burden. Measuring earnings would require linking these 1098T data with earnings data from tax returns at some point in the future. But note that this linkage would occur within one agency (the IRS) and not entail any data sharing. Perhaps the most important strength of the 1098T approach is that it is a “minor change to an existing process” and “does not entail repealing a legislative prohibition on a unit-record system that is endorsed by certain stubborn constituencies.”

Reformers should also work to facilitate state access to earnings data.
beyond their own borders. The federal government should immediately help states obtain IRS tax data (with the appropriate concerns for protecting privacy). Whether the IRS has the statutory authority to share these data with the states is contentious, and several requests from states have been shuffled off to bureaucratic never-never land. There is interest in using the U.S. Census Longitudinal Employer-Household Dynamics data, although the ultimately successful negotiations between LEHD and the University of Texas system took years to complete.²²⁹

A weaker alternative is to invigorate the existing Wage Record Interchange System (WRIS 2).²³⁰ WRIS 2 is a voluntary consortium of states that have agreed to answer data queries from other member states. Forty-three states, the District of Columbia, and Puerto Rico participate in WRIS 2.²³¹ It works like this: A member state needs a particular set of data, but has not found a set of students in its own state Unemployment Insurance (UI) wage database. It submits a list of the Social Security numbers of these missing students to the WRIS 2 consortium clearinghouse. The members of the consortium then run these records against their own UI data and return to the WRIS 2 clearinghouse earnings data on the records they have matched. The clearinghouse combines all the data received from member states and returns the earnings data to the requesting state. This is designed to overcome the interstate mobility of Americans, who may not live or work in the state in which they were educated and hence would not be found in their “home” state’s UI database. Obviously, this is a rather circuitous and burdensome route — and many states in the consortium have reported disappointingly low match rates.

In the meantime, efforts continue to try to improve and increase use of the WRIS 2 system. The number of states in the consortium continues to grow. Recently, the state of Iowa’s Workforce Development office has offered to submit the records of any school in Iowa to the WRIS 2 system, widely advertising and systematizing efforts to increase the use of WRIS 2. But, again, this is not the most efficient way to tap into the earnings data of individuals who work in states other than the one where they attended school.

REVOLUTIONARY REFORMS

A new administration could take a far more aggressive approach, replacing current policies that don’t work with streamlined federal efforts to collect data for the benefit of both students and taxpayers.
Such a plan would start by using federal policy to foster state-level data-collection efforts.

States invest large amounts of money in their post-secondary systems because post-secondary education is viewed as a human-capital investment that will help the state remain economically competitive. At the current time, states also “own” the student-level data (often built with support from federal-state longitudinal data system, or SLDS, monies). In addition, even as the federal government has expanded its role in education, state and local governments retain the lion’s share of legal and regulatory power over post-secondary education. So even while the federal government’s share of dollars flowing into colleges and universities is larger than the state and local investment, most of the federal dollars come through Title IV student-aid programs—a blunt instrument for effecting change. In contrast, most post-secondary students are enrolled in public colleges and universities, meaning that cultivating state partnerships is essential to future reforms.

But what should those partnerships look like? One option would be to condition the flow of federal money on states meeting certain standards or engaging in certain activities. The federal government, for example, gave hundreds of millions of dollars to states to build longitudinal data systems.²³² From the program’s inception until almost all the money in the grant had been spent, the federal government did not have a “use requirement” as a condition of the award. As a result, hundreds of millions of dollars were spent on building data warehouses that allowed very little public access. This is an example of what the federal government should not do. But it also contains a lesson that could direct future funding. Any new funding to support state data work must include a use requirement to make the data available to the public (again, with privacy protections in place).

Partnerships could also be built around more efficient ways of measuring the earnings outcomes of students. As noted above, the WRIS 2 system allows states to access earnings data for students no longer working in the state. However, the WRIS 2 consortium is an inefficient method for doing this. The federal government could facilitate states’ access to IRS tax data by entering into cooperative agreements with state agencies.

This approach offers several advantages. State data systems already record students’ programs of study—and the matched data would be at the student level rather than the institution level used in the Scorecard.
Further, state data systems already encompass students without federal aid. And finally, state data systems usually cover far earlier cohorts of students than any data held by the federal government. All of this would allow far deeper, far longer, and far more meaningful reporting of wage outcomes than the federal government will be able to provide any time in the near future. (One downside: most state systems do not contain information about students in private schools—but that is slowly changing, and that trend could be accelerated with federal incentives.)

Such cooperative agreements would allow states to tailor these wage-outcome data for their own policy purposes (including, for example, performance-based budgeting), would cut several years off the federal timetable for gathering program-level data, and could cover all students, especially those in public institutions, rather than just Title IV students.

Ideally, policymakers would replace current federal data-collection efforts with a federal data system capturing student-level information. This is the single most important change and the one that is least likely to happen. This idea has gone by many names: SUR (Student Unit Records), SURS (Student Unit Record Systems), and SURDS (Student Unit Record Data Systems). Not only are these acronyms ungainly, but they are also guaranteed to set off political debates, falling along all too predictable ideological lines (see the response to the Department of Education’s efforts to allow researchers to access FSA data, noted earlier). There seems to be no easy path forward. The Know Before You Go Act introduced in 2015 by Senators Wyden, Rubio, and Warner seems like a reasonable strategy but has received no more traction than the 2013 Wyden-Rubio proposal.²³³

The most aggressive actions to spread consumer-oriented information may no longer fall exclusively under the jurisdiction of the Education Department. Most notably, the Consumer Financial Protection Bureau is taking aim at many practices in higher education, including the absence of data. For example, the CFPB, under a “Know Before You Owe” rubric, has issued a financial-aid shopping sheet that gives students guidance about how much they could reasonably borrow, and helps students identify schools that use the shopping sheet (while castigating those schools that do not).²³⁴ And the Obama administration has pushed new data strategies forward through executive action (see especially the College Scorecard).

However, a new, more conservative administration would not likely pursue these executive branch-based approaches. The way past the
current deadlock over data seems to be two-fold. First, more states need to be brought into the national policy discussion; states are already leading the way in developing detailed student-level data systems that merge multiple existing data systems to provide information to students about the costs and rewards of their college choices. Whether working individually or in consortia, states should be supported and encouraged in these efforts. And innovative states, such as Tennessee and Texas, should have the support of the next administration as they try to crack open the bottleneck on wage data.

Second, somehow, the concerns for privacy and data security need to be addressed head-on. All too often these concerns are used by entrenched interests in higher education to protect themselves from inconvenient truths about costs, debt, and rewards. So long as the federal government suffers humiliating data breaches, opponents of better data systems built on the bedrock of student-level data will win, and students and taxpayers will be stuck with unmanageable debt and schools that don’t get their students across the finish line.
Policy Reforms to Strengthen Higher Education


185. See 34 CFR § 602.16.1, “Accreditation and Preaccreditation Standards,” www.law.cornell.edu/cfr/text/34/602.16: “The agency’s accreditation standards effectively address the quality of the institution or program in the following areas… (v) Fiscal and administrative capacity as appropriate to the specified scale of operations.”

186. Ibid.: “The agency’s accreditation standards effectively address the quality of the institution or program in the following areas… (x) Record of compliance with the institution’s program responsibilities under Title IV of the Act, based on the most recent student loan default rate data provided by the Secretary, the results of financial or compliance audits, program reviews, and any other information that the Secretary may provide to the agency.”

187. Ibid.: “The agency’s accreditation standards effectively address the quality of the institution or program in the following areas… (vii) Recruiting and admissions practices, academic calendars, catalogs, publications, grading, and advertising… (ix) Record of student complaints received by, or available to, the agency.”


188. This extremely narrow window is slowly being opened.

189. For additional information, see College Measures, “College Measures: Improving Higher Education Outcomes in the United States,” www.air.org/center/college-measures.


193. Application programming interfaces (APIs) define how different pieces of software can readily access data held, for example, in a data library. This allows, for example, the data released along with the federal government’s own Scorecard website to be repurposed by others.


195. The specter of a testing regime for colleges and universities that would immediately be compared to the mandatory tests of No Child Left Behind should alone be enough to give the government pause.


208. Ibid.


213. See the various reports and databases at College Measures, www.air.org/center/college-measures.


217. Email communication from former Department of Education staffer Ben Miller, June 2016. For more detail, see Miller’s testimony: Ben Miller, Testimony to U.S. House of Representatives Education and the Workforce Committee.

218. For instance, one recent paper suggested that experts in FSA’s new Data Office could make themselves available to respond to ad hoc research requests and/or solicit a Request for Information that would allow outside researchers and policymakers to suggest topics of interest. See Matthew Soldner and Colleen Campbell, *Using—and Improving—Federal Student Aid Data Systems to Support Policy Analysis*, Institute for Higher Education Policy, May 2016, www.ihep.org/sites/default/files/uploads/postsecdata/docs/resources/using_and_improving_fsa_data_systems.pdf.


225. Perhaps not surprisingly, this announcement was not greeted warmly by everyone. Privacy concerns can be overstated to thwart many needed efforts to improve performance, but these concerns are real, especially given major data breaches involving government data. For example, see Jane Robbins, “The Feds Call It ‘Research.’ We Call It Violating Child Privacy,” *Townhall*, May 6, 2016, http://townhall.com/columnists/janerobbins/2016/05/06/the-feds-call-it-research-we-call-it-violating-child-privacy-n2158642.

Policy Reforms to Strengthen Higher Education

Systems to Support Policy Analysis.
228. Whitehurst and Chingos, Deconstructing and Reconstructing the College Scorecard.
230. The original WRIS system was focused on tracking the earnings outcomes of participants in workforce training programs. WRIS 2 opened up the system to educational institutions. See U.S. Department of Labor, “Wage Record Interchange System 2,” https://doleta.gov/performance/wris_2.cfm.

PROMOTING INNOVATION IN HIGHER EDUCATION
235. The views expressed here are those of the authors and do not reflect those of their employers.

175