DOCKET ID ED-2022-OUS-0140

My name is Robert Kelchen and I am a higher education researcher who studies higher education finance and accountability. I am also a professor of education at the University of Tennessee, Knoxville and the data editor of *Washington Monthly* magazine's annual college guide and rankings. All opinions in this document are my own and do not necessarily represent the views of any of my affiliated organizations or collaborators.

My comments below cover each of the seven key questions that ED posed in its request for information.

**Question 1: What program-level data and metrics would be most helpful to students to understand the financial (and other) consequences of attending a program?**

Prospective and current students generally want to know what it will take for them to get through their desired program of study (assuming they know what that is upon entering the institution) and what the outcomes are for students after leaving college. This includes several different data points:

(1) Program-level completion rates. Right now, the only data on completion rates are at the institutional level, and only for undergraduate students. But whenever students are admitted to particular programs, it is fairly straightforward to create completion rates. Students are typically directly admitted to particular programs at the graduate level and for vocationally focused undergraduate programs. The challenge comes with so-called "traditional" undergraduate institutions at the two-year and four-year levels, where students are frequently admitted without majors and change majors on a regular basis. As a result, it may make sense to show program-level completion rates for undergraduate certificate programs and all graduate programs while falling back on institutional completion rates for associate and bachelor's degree programs.

Another concern with program-level completion rates is that institutions could try to change admissions processes to not assign students a major at entry. It may be possible for programs to provide input to a centralized admissions office on who to admit and place caps on the number of admitted students while technically not admitting to a particular program. This is frequently done among bachelor's degree-granting institutions, and it could be expanded to graduate or certificate programs to try to game the system. It is unclear whether colleges would try to take that step, but it bears watching.

(2) Time/credits to completion and sticker/net prices paid to complete. As different programs at the same institution have different requirements, tuition rates, and student charges, there can be large differences in what students can expect to invest in terms of time and money based on their chosen program of study. Differential tuition policies, in which students are charged a premium to major in high-cost or high-value fields, are common at public research universities and can

affect students' eventual majors.[1] Additionally, fields such as education and engineering have historically required more credit hours than the norm, increasing student charges.[2]

For example, the accountability webpage for Texas A&M University shows that students in engineering take approximately 4.4 years to complete bachelor's degrees compared to 3.8 years in liberal arts.[3] Engineering students pay an additional $1,000 or so more per semester than liberal arts students in differential tuition, and this is before costs of textbooks and laboratory supplies.[4] My quick back-of-the-envelope calculation suggests that engineering majors at the College Station campus pay at least $14,000 more in tuition and fees (before any discounts) than liberal arts majors.

(3) Student debt, ideally separated by completion status if possible. The current program-level College Scorecard already has reasonable metrics to emulate, although I struggle with how to treat Parent PLUS loans. Should they be combined with student debt, or are they truly separate?

(4) Earnings after leaving college, again by completion status if possible. Again, the current program-level Scorecard is a good guide. The question then becomes when earnings should be reported. From a back-end perspective, it is easy for ED to provide a dataset with data from throughout former students' time in the labor market. But for a public-facing site, decisions have to be made about what time period to feature. My gut instinct would be to feature earnings from roughly 3-4 years after leaving college, as this would allow individuals time to get established in the labor market and allow for special cases such as fellowships and medical residencies.

Another question to wrestle with for program-level earnings is how to handle students who go on to get an additional degree. There are substantial variations by major in the share of students who pursue further education. For example, students who graduate with a bachelor's degree in science are nearly twice as likely to pursue a graduate degree as business majors.[5] For the majority of science majors planning on earning another credential, seeing the earnings of students with just a bachelor's degree may not be relevant. It also raises cell size issues.

**Question 2: What program-level data and metrics would be most helpful to understand whether public investments in the program are worthwhile? What data might be collected uniformly across all students who attend a program that would help assess the nonfinancial value created by the program?**

First of all, I consider the data of interest to students to also be of interest to the public/taxpayers. But there are several other metrics that would be of interest in determining the value of a program to society, and here I focus on the federal government as the relevant stakeholder instead of the state government.

---

[1] Stange, J. (2015). Differential pricing in undergraduate education: Effects on degree production by field. *Journal of Policy Analysis and Management, 34*(1), 107-135.
[2] https://nces.ed.gov/programs/digest/d07/tables/dt07_314.asp.
[3] https://accountability.tamu.edu/All-Metrics/Mixed-Metrics/Time-to-Degree.
[4] https://tuition.tamu.edu/.
[5] Baum, S., & Steele, P. (2017). *Who goes to graduate school and who succeeds?* Urban Institute.

(1) Federal income taxes paid by former students. This would help to construct a measure of return on taxpayer investment and the federal government already has the data. It would also be helpful to recognize federal income taxes paid by for-profit colleges here in some way, as these institutions rightly note that they do return some funds to the federal government through tax payments. It would also be great to get data on public benefits used, but that may be too heavy of an interagency cooperation lift to work out.

(2) More information on income-driven repayment subsidies and loans that were defaulted upon. The amount of federal student loan dollars in income-driven repayment was rising quickly before the 2020 repayment pause, and that amount will only accelerate due to the three-plus year repayment pause and the Biden administration's plan to make income-driven repayment plans more generous to borrowers.

(3) A clearer definition about what should count as "public service." Right now, the definition of working for most government and nonprofit agencies is fairly straightforward to use to determine whether someone is eligible for Public Service Loan Forgiveness. But this definition splits programs of study in ways that may not be intended. For example, a nurse working for a nonprofit hospital chain is currently eligible for PSLF while a nurse for a for-profit hospital chain is ineligible. It is time to rethink whether all graduates in certain majors should be considered as public service, or if public service should be defined by the employment classification (Standard Occupational Classification). This could add to the current PSLF eligibility matrix, or replace it for this purpose.

**Question 3: In addition to the measures or metrics used to determine whether a program is placed on the low-financial-value program list, what other measures and metrics should be disclosed to improve the information provided by the list?**

There should be two main sets of metrics included in addition to program-level data. The first is information about any relevant warnings or sanctions levied by a member of the regulatory triad (federal government, state government, or accrediting agency). The federal government portion should include Heightened Cash Monitoring, financial responsibility, letters of credit, and any other relevant warning. The state government portion should include any relevant warnings or investigations, although this requires either states to cooperate or institutions to voluntarily disclose concerns. The accreditor portion should include any warnings at the institution level or the program level if a programmatic accreditor is relevant in that case.

The second set of metrics is a general set of outcomes for all students to account for the possibility of students switching programs of study. This is particularly relevant at the associate and bachelor's degree levels, as I previously discussed in this document.

**Question 4: The Department intends to use the 6-digit Classification of Instructional Program (CIP) code and the type of credential awarded to define programs at an institution. Should the Department publish information using the 4-digit CIP codes or some other type of aggregation in cases where we would not otherwise be able to report program data?**

The short answer to this question is yes, but this is an incredibly complicated area because a 6-digit CIP code is often not nuanced enough to answer important questions regarding value. I unpack some of these concerns below.

First of all, under Federal Student Aid's current reporting structure, program-level data are reported at the OPEID level instead of the UnitID level. If a program participation agreement with Federal Student Aid was set up at a system level, multiple institutions can have their data aggregated. In 2017, 20.2% of the 7,128 Title IV-participating institutions in IPEDS were so-called "child" institutions that were aggregated with parents.[6] This affects large public university systems such as Rutgers and Ohio State, and means that students at the main campus have their outcomes aggregated with the branch campus even though instructional resources and student characteristics may vary considerably.

Second, 6-digit CIP codes at a particular institution can include both online and in-person programs that are taught by different faculty and serve considerably different student bodies. A prominent example of this is the rise of online program management (OPM) companies, which perform administrative functions for institutions that wish to start and rapidly scale up online programs. OPM providers have received Congressional scrutiny and are of considerable policy interest.[7] Because of this, several OPM companies commissioned me to examine the outcomes of their former students using IPEDS and College Scorecard data. However, I was unable to do so because most online programs at these institutions also had sizable in-person programs in the same fields and the data were combined.[8]

I urge the Department to act on both of these issues, as they are likely to become more prominent in the future. I am particularly concerned about a trend in colleges rolling up accreditation and IPEDS data reporting to the system level, which happened in Indiana with Ivy Tech years ago and is underway in Maine, Vermont, and Connecticut among other states. For the purposes of reporting student outcomes, more thought needs to be given into what the institution level should be. But in the meantime, providing outcomes at the UnitID level and separating fully online and other (in-person or hybrid) programs would provide useful information for stakeholders.

With that being said, I strongly encourage rolling up data to the 4-digit CIP code when needed, and I would support making 4-digit CIP codes the default if that allowed for breakdowns by UnitID and/or online status. In the current data landscape, roughly 15% of all 4-digit CIP codes (at the UnitID level) have multiple 6-digit CIP codes, although that percentage has been slowly increasing in recent years.[9] If data were only reported at the 6-digit CIP level, institutions would have an incentive to split programs to try to stay below cell size requirements. Going to the 2-digit CIP level combines programs that often have substantially different outcomes, but that is

---

[6] Kelchen, R. (2019). Merging data to facilitate analyses. *New Directions for Institutional Research, 181*, 59-72.
[7] Vasquez, M. (2022, January 14). Online program management firms are thriving. And these Democrats want answers. *The Chronicle of Higher Education*. https://www.chronicle.com/article/online-program-management-firms-are-thriving-and-these-democrats-want-answers.
[8] Kelchen, R. (2022, March 29). Improving outcomes data for online programs. *Inside Higher Ed*. https://www.insidehighered.com/views/2022/03/29/how-improve-outcomes-data-online-programs-opinion.
[9] Blagg, K., Blom, E., Kelchen, R., & Chien, C. (2021). *The feasibility of program-level accountability for higher education.* Urban Institute.

less of a concern for providing consumer information than it would be for high-stakes accountability.[10] If needed, I am okay with providing institutional-level outcomes for that credential level if program-level outcomes are unavailable. But this needs to be appropriately flagged for data consumers.

**Question 5: Should the Department produce only a single low-financial-value program list, separate lists by credential level, or use some other breakdown, such as one for graduate and another for undergraduate programs?**

At the very least, it makes sense to break programs down by credential level because students often know whether they want to pursue, for example, a one-year certificate versus an associate degree. But I question whether credential level is sufficient to give students and policymakers the information that they need in order to make useful decisions. Unless lists are divided out by field of study in some way, the lists by credential level will be dominated by certain fields of study.

For example, I analyzed program-level earnings and debt data from the College Scorecard for a 2022 *Washington Monthly* feature on the best and worst undergraduate certificate programs in the ten most common fields of study.[11] The fifth-lowest HVAC maintenance program in the country had median earnings of $26,671 in 2019. Meanwhile, the highest-earning somatic bodywork program had earnings of $24,750 and only seven cosmetology programs in the country had earnings over $27,000.

Not separating out by field of study would repeat a key issue of the College Affordability and Transparency List—a recognition that some areas just have lousy outcomes. The public university portion of the list is dominated by a few states with historically high tuition rates, letting low performers in other states off the hook. As Obama and Biden Administration staffer Ben Miller once said about the list: "If you live in Pennsylvania, the lists tell you that you're basically screwed. If you don't live in Pennsylvania, the only thing the lists tell you is that you probably shouldn't move to Pennsylvania."[12]

If a student has their heart set on a certificate in somatic bodywork, a shame list by credential level will give them some useful information because the worst programs will be highlighted. But relatively low-performing HVAC programs will slide by when students could make better choices. So for an informational list, I recommend having filters by field of study and credential level.

**Question 6: What additional data could the Department collect that would substantially improve our ability to provide accurate data for the public to help understand the value being created by the program? Please comment on the value of the new metrics relative to the burden institutions would face in reporting information to the Department.**

---

[10] Blagg et al. (2021).
[11] Washington Monthly (2022). *2022 best colleges for vocational certificates.* https://washingtonmonthly.com/2022-college-guide/best-colleges-for-vocational-certificates/.
[12] Miller, B. (2013, July 1). No one's watching the watch lists. *New America.* https://web.archive.org/web/20130707112537/http://higheredwatch.newamerica.net:80/blogposts/2013/the_failed_tuition_watch_lists-86916.

Some of the data elements that I discussed will come from federal sources, and I do want to highlight the upcoming ability to look at outcomes by race thanks to adding a question about race to the FAFSA. But there are a few sources that will need to come from institutions, and it is important to discuss burden here.

The most substantial burden (barring finally lifting the provision on a federal student unit record data system, which looks unlikely in the next two years) is to get information on program-level completion rates. The burden should be manageable for graduate and certificate programs to which students are directly admitted, but it would add workload similar to reporting the number of credentials earned by program. To reduce burden, I would be fine with not requiring collection for associate and bachelor's degree programs. The value of data to prospective students and policymakers outweighs the burden. And if burden is a concern, I highly recommend cutting the rarely-used Academic Libraries survey as a trade-off.

Another sizable burden would be collecting better information about tuition charges and time to completion at the program level. This is the type of information that students and families need to know as they plan out how to afford college, and it is directly relevant to policymakers trying to encourage enrollment in high-cost fields such as engineering and nursing. Tuition charges for selected programs are in IPEDS at the graduate level, and it is at the very least important to add better information about differential tuition and fee charges for the most common programs.

A more modest upfront burden would be separating out exclusively fully online and other (in-person and/or hybrid) programs to provide better information about value. IPEDS has indicators at the 6-digit CIP level for whether a program can be completed via distance education, but what is really needed is an indicator for whether a program can *only* be completed via distance education. This would represent an initial investment of time to first implement, but the ongoing burden should be minimal.

**Question 7: What are the best ways to make sure that institutions and students are aware of this information?**

At this point in time, I think that many institutions are already aware of their programs' performance. Program-level data on debt and earnings have been available in the for-profit sector for a number of years thanks to efforts to implement gainful employment regulations. My research on the 2017 gainful employment data release suggests that for-profit colleges were more likely to shut low-performing programs, even though it was clear by the time of the data release that gainful employment was not likely to be implemented for the next several years.[13] And for programs not covered under gainful employment, the program-level College Scorecard data release provided useful information about student outcomes that was often not available to institutional leaders.

However, the initial effect of program-level debt and earnings information has likely worn off. It could be possible that leaders would respond to new data elements, such as separating the

[13] Kelchen, R., & Liu, Z. (2022). Did gainful employment regulations result in college and program closures? *Education Finance and Policy, 17*(3), 454-478.

outcomes of fully online and in-person/hybrid students. But the record of institutions making changes based on information that they already knew is less optimistic. The College Affordability and Transparency List is a good example. Research by Dominique Baker of Southern Methodist University found that being included on the tuition shame list had no effect on future tuition, as colleges with high tuition already knew that they had high tuition and that it was often out of their control.[14] Additionally, 82% of colleges on the 2019 tuition shame list said that they had no plans to reduce costs that drove higher tuition charges, suggesting that institutions are not looking to change their actions simply by being shamed.[15]

I do not expect these lists to have large direct effects on students. Research on the College Scorecard's high-profile introduction found very small effects on student application patterns, and those were driven by more affluent students who were already applying to multiple colleges.[16] Another working paper found a very small change in Google keyword searches for colleges with strong outcomes.[17] Additionally, Baker's research on the College Affordability and Transparency List found no effects on enrollment.[18]

If these lists are effective, it will require other stakeholders, such as college access organizations, guidance counselors, and journalists, to use the information. But this has already been happening with the College Scorecard, and the effects are at most modest.

I appreciate having the opportunity to provide feedback on these seven questions and please do not hesitate to contact me (rkelchen@utk.edu) if you have any questions.

Sincerely,

Robert Kelchen

[14] Baker, D. J. (2020). "Name and shame": An effective strategy for college tuition accountability? *Educational Evaluation and Policy Analysis, 42*(3), 393-416.
[15] https://collegecost.ed.gov/affordability.
[16] Hurwitz, M., & Smith, J. (2018). Student responsiveness to earnings data in the College Scorecard. *Economic Inquiry, 56*(2), 1220-1243.
[17] Huntington-Klein, N. (2017, April 25). *The search: The effect of the College Scorecard on interest in colleges.* Working paper. https://www.nickchk.com/Huntington-Klein_2017_The_Search.pdf.
[18] Baker (2020).