The Obama Administration’s proposed Postsecondary Institutions Ratings System (PIRS), announced last August, is a reaction to frustrations about the rising cost of college. Although funds for the federal Pell Grant more than doubled between the 2007-08 and 2011-12 academic years, the net price of attendance at four-year public universities increased by 10 percent during the same period. PIRS will rate colleges in three main areas, possibly including some of the following measures:

- **Access**: Percentage of enrolled students receiving Pell Grants, effort in enrolling first generation students
- **Affordability**: Net price of attendance, tuition levels, student loan burdens
- **Outcomes**: Retention and graduation rates, employment, wages

The Department of Education (ED) conducted public meetings to collect ideas about potential measures and mechanisms and received comments from stakeholders through the request for information process. While some comments have been supportive, the vast majority have been

**Abstract**

This policy brief describes potential structures for the Obama Administration’s proposed Postsecondary Institutions Ratings System (PIRS), which is expected to rate colleges on measures of access, affordability, and outcomes by 2015-16. The brief discusses key issues and considerations for selecting measures and designing an effective system. It concludes with a set of specific recommendations.

**About The Author**

Robert Kelchen is an assistant professor in the Department of Education Leadership, Management and Policy at Seton Hall University. He is also the consulting methodologist for *Washington Monthly* magazine’s annual college rankings. All opinions in this piece are the author’s own and do not necessarily reflect the opinions of WISCAPE, Seton Hall University, or *Washington Monthly*. Contact the author at robert.kelchen@shu.edu.
opposed to federal ratings. The Department then held a technical symposium on February 6 with about 20 presenters (myself included) discussing various aspects of PIRS. A summary of the symposium in The Chronicle of Higher Education notes a healthy amount of skepticism with the idea of ratings.

After collecting feedback, ED plans to move quickly in order to meet an ambitious timeline. The goal is to have draft ratings ready this spring, with a final set of ratings in place for the 2015-16 academic year and to tie the ratings to Title IV aid eligibility by fall 2018. Doubts remain that ratings will be tied to aid, however, as this would require the approval of Congress and the next president.

If federal ratings are unlikely to be tied to financial aid, are they still worth doing? The federal government should consider college ratings if the expected benefits outweigh both the expected costs of compiling the ratings and the next best use of those funds. Whether the ratings will be cost-effective remains to be seen, but research on K-12 accountability policy suggests that ratings and incentives can have a substantial impact on school outcomes.

Whether the ratings will be cost effective remains to be seen, but research on K-12 accountability policy suggests that ratings and incentives can have a substantial impact on school outcomes.

Access
The most commonly referenced measure of access is the percentage of students receiving Pell Grants. Since about 85 percent of students receiving Pell Grants have a family income below $40,000 per year, it is in some respects a good measure of low-income status. However, research by Mark Kantrowitz suggests that 2.3 million additional students would have qualified for a Pell Grant had they filed the FAFSA; this would represent a 25 percent increase in the number of Pell recipients. Community colleges have noted in their public comments about PIRS that using Pell Grants as a measure of access could understate the percentage of needy students at many colleges.

Affordability
The net price of attendance (defined as the total cost of attendance less all need-based and merit-based grant aid) will almost certainly be included in PIRS. The federal government currently calculates net price for all first-time, Postsecondary Education Data System (IPEDS). This system collects select information primarily on first-time, full-time students. Although policy analysts and institutional associations have noted the limitations of IPEDS data, it is currently the best national data source available. This section focuses on key measures of access, affordability, and outcomes that are readily available through IPEDS or other sources, discussing key issues to consider while weighing their inclusion in the federal ratings system.

The remainder of this policy brief draws upon my work as the methodologist for Washington Monthly’s annual college rankings and research on accountability policy to describe potential structures for PIRS. The brief concludes with a discussion of key issues that must be addressed while constructing the ratings system and a set of specific recommendations.

Key Ratings Components
To develop ratings by fall 2015, ED will have to rely on institutional-level data from the Integrated Postsecondary Education Data System (IPEDS). This system collects select information primarily on first-time, full-time students. Although policy analysts and institutional associations have noted the limitations of IPEDS data, it is currently the best national data source available. This section focuses on key measures of access, affordability, and outcomes that are readily available through IPEDS or other sources, discussing key issues to consider while weighing their inclusion in the federal ratings system.
full-time students who receive any financial aid as well as for five income brackets (ranging from $0-$30,000 to above $110,000 per year). This focus on first-time, full-time students provides an incentive for colleges to allocate financial aid to students who are right out of high school instead of to returning adults. An alternative that would not penalize part-time or returning students would calculate the average net price for families with household incomes of $75,000 or less, as in Washington Monthly’s 2013 list of “best bang for the buck” colleges.

Student debt upon leaving college can be a better measure of affordability than net price, because student debt is partially driven by a family’s ability to pay. However, colleges with small endowments and large numbers of low-income students may be at a disadvantage without adjusting for student characteristics. It is also essential to include both dropouts and graduates in this measure to better reflect all students’ outcomes. Finally, data quality is an issue as some university systems capture loan data at the system level instead of for individual institutions, and transfer students may hold loans from attending multiple institutions.

Outcomes
The most-discussed outcome measure is graduation rates, which are currently collected for first-time, full-time students who stay at the same institution. For selective four-year institutions, the IPEDS completion measure does a reasonable job following a student cohort, but the data collected often understate the true success rates of students at less-selective four-year institutions and community colleges with high transfer rates. Data from the National Student Clearinghouse (NSC) show how a substantial number of students transfer from community colleges to four-year colleges without earning an associate degree; while these student are not counted as a success in IPEDS, they should be in PIRS. A potential short-term solution is adopting the Student Achievement Measure (SAM), a joint project of many higher education membership organizations, which uses NSC data to track students across institutions.

There has been a push in policy circles to report graduation rates for more subgroups of students. While IPEDS already contains graduation rates by race/ethnicity and gender, an important subgroup is missing: Pell Grant recipients. Colleges currently are required to collect Pell graduation rates under the Higher Education Act, but they are not required to submit the data to IPEDS. As a result, the only complete data source for Pell graduation rates is U.S. News and World Report [collected through the Common Data Set], but much of the data are only available by subscription. Graduation rates for other student subgroups could also be included in PIRS, as could measures of closing gaps in completion by race/ethnicity, gender, or family income.

Student loan default rates reflect both earnings and initial borrowing amounts. There are three primary issues with using default rates as a measure. First, not all loans are tracked (PLUS loans are excluded) and the longest default window currently tracked is three years into the standard, ten-year repayment period. Additionally, about one million students attend community colleges that do not participate in federal loan programs, and data on default rates is not available for these institutions. Finally, the rapid growth in the number of students enrolled in income-based repayment (IBR) programs could render default rates useless, as payments drop to zero if a student’s income is sufficiently low.

Some measure of employment or wages is widely expected to be included in PIRS, especially as ED has created space for employment outcomes on the College Scorecard, although the data are not yet available. Not surprisingly, most of the higher education community is strongly opposed to the principle of measuring colleges by earnings. Some argue that the purpose of higher education is not strictly vocational, while others in lower-wage fields such as education and social work contend that their missions are more valuable to the public than their graduates’ average wages suggest.
If wages are to be used, a few guiding principles should be considered. First, like student loan volume or default rate measures, the measures should include dropouts as well as graduates. Second, the data must not come from a convenience sample of satisfied alumni, such as that compiled by Payscale. Instead, the data should come from either a statistically representative sample of alumni (in the shorter term) or from state and/or federal unemployment insurance records (in the longer term). An analytic concern is whether to adjust for the local cost of living. $50,000 per year in the New York City metropolitan area has far less purchasing power than $50,000 per year in the rural Midwest. While making such an adjustment may be necessary, it could pit colleges in different parts of the country against each other while raising concerns about the proper cost multiplier.

While using salary figures in PIRS raises many issues, the percentage of former students earning a wage over a certain threshold could be a useful alternative measure. For instance, a successful outcome could be defined as having a salary above 200 percent of the poverty line for a family of three ($39,060 in 2013). Although the earning threshold should be discussed in more detail, it is important to set the cutoff so people pursuing careers in lower-wage professional fields can qualify when working nearly full-time.

Cost-effectiveness measures could also be included in PIRS, combining student outcomes with the resources to generate those outcomes. Some research has examined cost-effectiveness with respect to graduation rates, using measures such as the total government subsidy after taxes and education and related expenditures to look at the cost per degree or certificate completed. Another option is to use adjusted measures, such as the cost-adjusted graduation rate performance measure introduced in the 2012 Washington Monthly rankings.

Key ratings decisions

In order to design a fair and effective ratings system that can accurately identify both high-performing and low-performing colleges, the Department of Education has a series of important decisions to make over the next few months. These are outlined below.

Question 1: What should the scoring system look like?

The way in which the scoring system is constructed for PIRS will have significant implications for colleges, the general public, and the federal government. A scoring system must be easily understood by all stakeholders, transparent, and viewed as fair by all parties involved. Some potential scoring options include:

A-F grades
This type of ratings system is used in some states’ K-12 accountability systems. Based on those experiences, this system may be difficult politically and is likely to result in tension with colleges.

Pass/fail grades
A pass/fail system would be similar to how the accreditation process for financial aid eligibility currently works, but it does not provide the level of gradation desired for PIRS.

3-5 unique ratings categories
This is the most likely outcome and could take many forms. An example is to use gold, silver, bronze, and lead ratings as proposed by David Bergeron at the Center for American Progress. Others have proposed to use a system of color ratings, such as red/yellow green.

It is critical that ED use broad ratings categories in PIRS to avoid creating an underlying set of college rankings. Colleges will oppose the Administration’s efforts if they envision their individual rankings being published. One option to reduce these concerns is to use categorical
metrics. For example, graduation rates could be divided into three categories: low (below 50 percent), medium (50-75 percent), and high (above 75 percent).

**Question 2: How should colleges be classified for the ratings?**

Many people involved in the ratings discussions have urged ED to put colleges into narrowly defined peer or comparison groups with similar institutions. While colleges may see this approach as being the most fair, there are concerns over how comparison groups are constructed. Most college rankings use the Carnegie classification system to group colleges by their educational mission and highest degree offered. This classification system make sense for measures that take research activity into account (a key factor in the Carnegie system), but the distinctions matter far less for more basic measures like those discussed to this point for PIRS.

A better option is to use broader classifications, rating all four-year institutions together and then dividing two-year colleges into two groups based on degree offerings. Although there are many differences among four-year institutions, nearly all grant bachelor’s degrees at the undergraduate level. Two-year colleges should be grouped based on whether more students are enrolled in associate-granting programs or certificate-granting programs, as there are key program length and mission differences between these two groups.

**Question 3: Should measures be adjusted for student and/or institutional characteristics?**

One of the primary arguments for using narrow comparison groups in PIRS is that the types of students and financial resources available to an institution (its inputs) would be far more similar than in a broad comparison by sector. But since a ratings system requires at least some comparisons to be made across dissimilar colleges, the question then becomes how to adjust for institutional differences in inputs that can affect outputs such as graduation rates and wages. Input-adjusted or risk-adjusted metrics use regression techniques to control for characteristics such as the percentage of students receiving Pell Grants, race/ethnicity, gender, and academic preparation. This generates a predicted value of an outcome, which is then compared to the actual value to estimate institutional value-added.

Input-adjusted metrics are controversial because they allow colleges serving large percentages of disadvantaged students to have lower success rates while still meeting their goals. For example, a college serving 75 percent Pell Grant recipients may be expected to have a graduation rate of 20 percent. If that college than graduates 30 percent of its students, should it be considered successful? There is also debate over what inputs should be included in the regression model, although some think that educational expenditures should be excluded to avoid giving more expensive colleges an advantage over institutions that are more affordable.

One possibility to alleviate concerns about input-adjusted metrics is to require that colleges meet minimum criteria for both absolute and input-adjusted performance. For instance, to be included in *Washington Monthly*’s list of “best bang for the buck” colleges in its 2013 College Guide, colleges had to have a graduation rate of at least 50 percent and at least equal to their predicted graduation rate. This method has limitations for colleges with extremely high graduation rates, but these could be resolved by loosening or waiving the input-adjusted requirement.

**Question 4: How should individual metrics be weighted?**

Value judgments will determine what metrics go into the ratings, along with their respective weights. The most straightforward approach is to assign equal weights to access, affordability, and outcomes and to equally weight all outcomes
within each of the three metrics, but this is also a value judgment. Whatever the weights are, it is imperative that the results be robust across a range of reasonable model specifications. Colleges that are placed in the lowest and highest rating categories under equal weights should also be in the same category if one of the three categories is given a weight of one-half instead of one-third. If results are sensitive to small changes in weights, sensitivity analyses should be considered to better reflect uncertainty.

Recommendations

What follows are recommendations to increase the potential effectiveness of a federal ratings system.

1. Continue to reach out to stakeholders.
   The Department of Education should continue to listen to students, faculty, colleges, and the general public. ED must give the higher education community a chance to comment on the draft ratings and be willing to adjust them based on feedback. This will help increase institutional buy-in for PIRS.

2. Focus on identifying the very worst and very best colleges on the chosen metrics.
   Given imperfect data and the goal to tie federal financial aid to ratings, ED should take steps to ensure that colleges placed in the lowest category are truly low quality. For instance, if the goal is to identify colleges in the bottom 20 percent on a given metric, it may be wise to instead identify the bottom 10 percent. This will result in fewer “false positives,” in which a college is falsely identified as unsuccessful.

3. Use multiple years of data when possible.
   Relying upon the most recent year of data may work well for large colleges, where large changes in reported outcomes are less likely to be a result of statistical noise. However, at smaller colleges the outcomes of just five or ten students could have a meaningful impact on reported graduation or student loan default rates. Using three years of data reduces the likelihood of random error.

4. Audit a sample of colleges’ data submissions for accuracy.
   Several colleges have submitted incorrect data to U.S. News and World Report recently, which has resulted in them being moved to the “unranked” category. Likewise, an incorrect data submission to IPEDS resulted in one college receiving a spot on Washington Monthly’s list of “best bang-for-the-buck colleges,” although it was removed after the error was discovered. To ensure accuracy, ED should randomly audit a small percentage of colleges submitting data each year and closely examine colleges that report large changes.

5. Be willing to adjust the ratings as more data become available, and keep pushing for better data.
   For the first few years, PIRS will have to rely on a relatively small number of measures. But the potential exists for more measures to be included in the future, including estimates of student learning and the satisfaction of former students. The higher education community and ED should encourage the collection of new data and consider including them in IPEDS going forward. Congress should also overturn its 2008 ban of a student unit-record (SUR) data system by passing the bipartisan Student Right to Know Before You Go Act, which would also make more data available to students, their families, and other interested parties.
Endnotes


3 Summaries of each of the four public forums can be found at http://www.ed.gov/college-affordability/additional-resources, and the full set of public comments is available at http://www.regulations.gov/#docketDetail.docketID=ED-2013-IES-0151.

4 My presentation is available at http://kelchenoneducation.files.wordpress.com/2014/02/kelchen_pirs_presentation.pdf.


10 For more information, see http://nces.ed.gov/ipeds/resource/institutional_net_price.asp.


13 For more information on the SAM, see http://www.studentachievementmeasure.org.


16 Data tables can be found at http://studentaid.ed.gov/about/data-center/student/portfolio.

17 For an example of the College Scorecard, see Seton Hall University’s page at http://collegecost.ed.gov/scorecard/UniversityProfile.aspx?org=s&id=186584.

18 The federal poverty guidelines are available at http://aspe.hhs.gov/poverty/13poverty.cfm#guidelines.


22 One potential concern is with community colleges that grant a small number of bachelor’s degrees, which are classified as four-year institutions in IPEDS but should be counted as two-year colleges in PIRS.


The Wisconsin Center for the Advancement of Postsecondary Education (WISCAPE) promotes the creation and sharing of ideas for addressing Wisconsin's postsecondary education challenges. The production and dissemination of publications are a major part of this effort.

WISCAPE Policy Briefs are succinct analyses that provide policymakers, practitioners, and others with knowledge and recommendations based on the latest research and best practices in the field.

Credits
Editing and Layout: Kari Dickinson and Rachel Nathanson

Send questions about WISCAPE publications to:
Kari Dickinson, Communications Manager, 608-265-6636, kadickinson@wisc.edu

Recommended citation for this publication: