

**To: Richard Reeves
National Center for Education Statistics
U.S. Department of Education
1990 K Street NW, 8th Floor
Washington DC 20006**

**ATTN: Postsecondary Institution Ratings System RFI
Docket ID ED-2013-IES-0151**

The Institute for Higher Education Policy (IHEP) submits these comments on the Administration's proposed Postsecondary Institution Ratings System (PIRS), published in the Federal Register on December 17, 2013. Celebrating its 20th year anniversary, IHEP is a non-profit, non-partisan research organization committed to promoting access to and success in higher education for all students, with a particular focus on populations that have been traditionally underserved by our postsecondary system.

IHEP applauds the President's proposal to create a system that would, first and foremost, provide better information to help students and parents choose colleges that offer them the "best value." Too many of today's students are paying far too much at institutions that offer them far too few chances for success. Ensuring that information is provided to students in a timely, easy-to-digest format (popularly referred to as "consumer information") can, in fact, act as a form of soft accountability, allowing students to "vote with their feet" in the direction of institutions that will serve them well.

However, given the immense – and growing – student and public investment in higher education, we cannot afford "soft accountability" alone. The federal government must use real stakes to improve – or sanction as needed – institutions that are not serving students well, which is why IHEP also supports the Administration's proposal to develop a rating system to assess and tie college performance to the distribution of Title IV financial aid.

That said, the design of the system will need to differ depending on its purpose: consumer information or institutional accountability. While most of the metrics might remain the same, they should be applied differently in different contexts. In short, IHEP recommends separate rating systems be developed for each purpose. In guiding the development of those systems, IHEP also offers the following three principles underlying our recommendations:

1. Critical information on post-college outcomes, completion, cost, and access is key
2. Contextualize information without undue complexity and in relation to the consequences
3. Collect better data, but don't delay providing information now

1. Critical information on post-college outcomes, completion, cost, and access is key

In our recommendations (which follow the format of the Request for Information), we offer four categories under which institutions should be measured in rating systems for both consumer and accountability purposes:

- Access: Who attends an institution?
- Progression and Completion: Who succeeds at an institution?
- Cost and Debt: How much do students invest in college?
- Repayment and Earnings: How do students fare after college?

We also outline the data availability and feasibility of specific metrics within each category, and how existing data could be improved for better use in the rating systems.

2. Contextualize information without undue complexity and in relation to the consequences

While most of the metrics provided to students might be similar to those used as accountability metrics, IHEP recommends that these be applied differently with respect to constructing ratings given their differing purposes. In fact, we recommend that the Department create multiple model rating systems – incorporating different metrics or different visual formats – and allow institutions, students, and other stakeholders to comment on them before deciding on a final system.

For consumer information purposes, we recommend that the Department consider the following guidelines when presenting ratings to students:

- Present and rate each metric without complexity, and do not (only) display a composite rating.
- Keep institutional groupings broad and do not adjust rating cutpoints on metrics.
- Complement the ratings with information on similar, but better-performing institutions.
- Invest in counseling and other supports that help students interpret and use ratings.

For accountability purposes, we recommend that the Department consider the following guidelines to create a system (or systems) that not only protect students and the public investment, but also support institutions to improve outcomes:

- Be clear about purpose and consequences in designing the system(s).
- Consider complexity, but strive for simplicity.
- Set low, unadjusted performance floors for Title IV eligibility.
- Apply graduated eligibility using more nuanced mechanisms (such as a performance matrix), but avoid composite indicators.
- Consider a positively-framed incentive-based system.

There is a lot of debate about whether the Department should only compare institutions to other similar institutions or should adjust for student characteristics. We strongly feel that students need unvarnished information about college costs and completion and there is no justification for adjusting those measures in a consumer context. Making such adjustments may warrant consideration in an accountability context, although we suggest caution for at least two reasons. First, many incorrectly interpret adjusted outcomes as a “ceiling,” thereby lowering expectations for students and institutions. Second, “expected” or “adjusted” outcomes are calculated based on the current range of institutional performance. With fewer than half of students earning degrees on time, we should take care not to reinforce the status quo in the rating systems.

3. Collect better data, but don't delay providing information now

Good data are clearly critical to well-designed consumer information and accountability systems. However, the absence of perfect data should not be used as an excuse to avoid providing any information. Currently available data are undoubtedly strong enough to at least identify the worst-of-the-worst institutions from both consumer and accountability perspectives. As the Department moves forward with developing the rating systems, it also should work to collect higher quality postsecondary data, including graduation rates for Pell Grant recipients, cumulative debt levels for completers and non-completers, upward and lateral transfer rates, and workforce outcomes for graduates, which we discuss in detail herein. The Department should also continue to integrate or link with data systems operated by Federal Student Aid and/or federal agencies outside of the Department of Education to better measure student outcomes, particularly after college.

We thank you for the opportunity to share our suggestions on how to design the college rating systems and look forward to working with the Department and the Administration further on these efforts.

Sincerely,



Michelle A. Cooper, Ph.D.
President
Institute for Higher Education Policy



Jennifer Engle, Ph.D.
Vice President for Policy Research
Institute for Higher Education Policy

Table of Contents

- 1.0 Data Elements, Metrics, and Data Collection..... 5**
 - Possible Metrics for Rating the Performance of Postsecondary Institutions.....5
 - Access: Who attends an institution?.....5
 - Data Availability for Reporting: Access5
 - Potential PIRS Metrics: Access8
 - Progression and Completion: Who succeeds at an institution?.....9
 - Data Availability for Reporting: Progression & Completion.....9
 - Potential PIRS Metrics: Progression & Completion13
 - Cost and Debt: How much do students invest in college?15
 - Data Availability for Reporting: Cost & Debt15
 - Potential PIRS Metrics: Cost & Debt17
 - Repayment and Earnings: How do students fare after college?19
 - Data Availability for Reporting: Repayment & Earnings19
 - Potential PIRS Metrics: Repayment & Earnings24
- 2.0 & 3.0 Weighting and Scoring and the Development of Comparison Groups 27**
 - Rating System for Students27
 - Present and rate individual measures, but not a composite rating.....28
 - Keep institutional groupings broad and do not adjust rating cutpoints on metrics31
 - Complement the ratings with information on similar but better-performing institutions33
 - Invest in counseling: The best consumer information system35
 - Rating System for Accountability35
 - Be clear about purpose and consequences36
 - Allow for complexity, but strive for simplicity36
 - Permit only minimal adjustments based on institutional mission.....36
 - Set a low, unadjusted performance floor for Title IV eligibility37
 - Apply graduated penalties using more nuanced mechanisms, but avoid composite indicators.....38
 - Consider a positively-framed incentive-based system45
- 4.0 Presentation of Ratings Information 46**
- 5.0 Existing Rating Systems 47**



1.0 Data Elements, Metrics, and Data Collection

Possible Metrics for Rating the Performance of Postsecondary Institutions

There are four broad categories and questions the Department should consider when choosing metrics for the rating systems. The four questions include:

- **Access: Who attends an institution?**
- **Progression and Completion: Who succeeds at an institution?**
- **Cost and Debt: How much do students invest in college?**
- **Repayment and Earnings: How do students fare after college?**

Within each of these categories are several core measures, most of which are applicable to both consumer information and accountability purposes, although these *measures* will need to be applied differently as *metrics* for these different purposes. Some of the data for the measures are readily available, some are available but could be improved (such as through disaggregation), and some are not currently collected at all. For each of these questions, we outline the core measures that should be made publicly available (if not currently), the availability of this information in IPEDS, and how the data could be improved if necessary. Ideally, each of these data points would be made available at the program-level, but this level of detail could be impractical if IPEDS remains the primary higher education data source, and if the ratings are to apply to all postsecondary institutions, as we recommend. Fine-grained, program-level information likely would require a more comprehensive data collection such as a student unit record system, which is currently banned at the federal level.

After identifying the measures that should be available, we outline possible metrics that could be used with existing data to rate the performance of postsecondary institutions, as well as how they might be improved before incorporating them into rating systems for students and policymakers.

Access: Who attends an institution?

The makeup of the student body provides critical contextual information to students – who want to know the demographic profile of the student body, policymakers – who need to understand which institutions are serving which populations, and institutions – as they set and measure progress towards access goals. Not only does this demographic information on college enrollments allow interested parties to measure status and progress on access-related metrics, it also provides context for understanding who does (and does not) succeed at institutions.

Data Availability for Reporting: Access

IPEDS data on college enrollments are fairly comprehensive. Data are available – in some form – by enrollment status (e.g. first-time, transfer), attendance pattern (e.g. full-time, part-time), degree-seeking status, race/ethnicity, gender, age, financial aid status, and, beginning in 2014-15, by military status. Table 1 details the availability and limitations of the core enrollment diversity measures, and

proposes improvements that would make the enrollment surveys even more useful and robust than they currently are in IPEDS for purposes of the rating systems.

While Table 1 includes a number of recommendations, one key improvement that could provide a more complete picture of enrollment diversity could be accomplished by simply shifting disaggregates from the fall enrollment survey to the 12-month enrollment survey. Currently, IPEDS reports both fall enrollment and 12-month enrollment counts. Fall enrollments can be cut in a variety of ways, including race/ethnicity, gender, age, degree/certificate-seeking status, attendance status, and enrollment status, while 12-month headcount enrollments are available only by race/ethnicity, gender, and an undergraduate/graduate distinction.

While the fall enrollment data are disaggregated more thoroughly, the 12-month headcount enrollment actually provides a more complete picture of enrollment at the institution because it captures students who enroll at times other than the fall, a measure that grows increasingly important in an era of 21st-century students. The undercounting of students in the fall enrollment survey is particularly problematic in community colleges and for-profits, which often admit and enroll new students throughout the calendar year. For instance in 2011-12, the 12-month enrollments at public community colleges include 3.6 million more students than the fall enrollment counts (10.6 million vs. 7.0 million), and at four-year for-profit institutions, 12-month enrollments are 1.6 times higher than fall counts (2.1 million vs. 1.3 million).¹ To obtain a more complete and accurate picture of student enrollment, the 12-month enrollment survey should add disaggregates for *at least* age, enrollment status, and attendance pattern. A truly comprehensive data system also would disaggregate on other key factors, such as dependency status (e.g. independent or dependent), disability status, language proficiency, and parental education. At this time, however, Table 1 focuses on a core subset of access-related measures, which were identified based on a review by IHEP of the major higher education data initiatives that have developed over the last five to ten years, such as Complete College America, Achieving the Dream, and Access to Success, among others.

¹ Institute for Higher Education Policy analysis of 2011-12 IPEDS 12-month and fall enrollment data.

**Table 1. Access: Who Attends an Institution?
Measures and Data Availability**

Measure/Data Element: % of freshman class and % of undergraduates by:	Data Availability: Are these measures currently collected in IPEDS?
Enrollment Status (e.g. first-time, transfer)	Partially Available: In IPEDS, Fall enrollments (but not 12-month enrollments) are disaggregated by first-time/transfer status. These data are currently available descriptively about the student body, not disaggregated for student outcomes. New IPEDS Outcome Measures will report completion by first-time and transfer status, but with limitations.
Attendance pattern (e.g. full-time, part-time)	Partially Available: In IPEDS, Fall enrollments (but not 12-month enrollments) are disaggregated by full- and part-time status. These data are currently available descriptively about the student body, not as a disaggregate for student outcomes. New IPEDS Outcome Measures will report completion by full-time and part-time status, but with limitations.
Degree-seeking Status	Partially Available: IPEDS disaggregates Fall enrollments (but not 12-month enrollments) by degree/certificate-seeking, non-degree/certificate-seeking, and degree/certificate-seeking status unknown. It does not disaggregate degree-seeking students from certificate-seeking students. The Graduation Rate Survey cohort disaggregates bachelor's-seeking students from other degree/certificate-seeking students, but does not disaggregate associate's degree from certificate-seeking students.
Major or Program of Study	Partially Available: IPEDS includes degrees awarded in each program, but not # or % of students enrolled in the program. Outcomes (e.g. completion rates) by program also not available.
Race/Ethnicity	Currently available in IPEDS
Gender	Currently available in IPEDS
Age	Partially available: IPEDS Fall enrollments (but not 12-month enrollments) are disaggregated by age categories (under 18, 18-19, 20-21, 22-24, 25-29, 30-34, 35-39, 40-49, 50-64, 65+, age unknown). These data are available descriptively about the student body, but not available as a disaggregate for student outcomes.
Income or Financial Aid Receipt	Partially available: IPEDS Student Financial Aid component currently reports % Pell among first-time, full-time freshmen and among all undergraduates, but not % receiving Subsidized Stafford loans and no Pell or % receiving neither Subsidized Stafford loans nor Pell. The Pell data are available descriptively about the student body, but not as a disaggregate for student outcomes. Income and financial aid receipt are also available from NSLDS, and potentially could be linked to outcomes, but for Title IV recipients only.
Military Status or Military Benefits Receipt	Partially available: Beginning in 2014-15, IPEDS will include the # of undergraduate and graduate students receiving Post-9/11 GI Bill benefits and the # receiving DoD Tuition Assistance. These data will not disaggregate members of the military from family members receiving the benefits, nor be available as a disaggregate for student outcomes.
Level of Academic Preparation	Partially available: No central source provides information on high school course-taking or remedial placement/course-taking. IPEDS includes 25th and 75th percentile SAT/ACT scores, but data are missing for many open access institutions, where remediation is most prevalent.

Potential PIRS Metrics: Access

Each of the ten access measures listed above provide critical information to both students and policymakers about who is being served by which postsecondary institutions. As such, we recommend the data on those measures be improved as noted above and made available to the public. However, for purposes of developing the rating systems, it is important to identify a few clear and compelling metrics for ease of presentation of the data in order to effectively impact consumer, institutional, and policymaker actions, the stated purpose of the systems.

Percent Receiving Pell Grants

To that end, we recommend including “**Percent Pell**” as a measure in the rating systems for both students and policymakers (although we caution the use of this metric to “excuse” or “lower expectations of institutional performance as we will discuss in more detail later). This metric provides a clear, well-targeted measure of the percent of the student body that is low-income, and ensures that institutions are not excluding low-income populations as a way to increase selectivity or performance on other metrics included in the rating systems. Additionally, low-income status is highly related to each of the other access measures above, establishing it as a reasonable proxy. As mentioned in Table 1, IPEDS currently collects the percent of Pell recipients at Title IV institutions, for first-time, full-time students and for all undergraduates. Using this measure in the systems would apply little to no additional reporting requirements on institutions, and Pell Grants are well-targeted at low-income populations (nearly three-fourths of Pell recipients have family incomes under \$30,000 annually²).

It should be noted, however, that this measure is a proxy for the percent of low-income population on campus, and is an imperfect measure of the overall income distribution since not all low-income students apply for or receive Pell Grants, particularly at low-cost institutions such as community colleges. This problem may be ameliorated, although probably not eliminated, since institutions may be incentivized to ensure that all Pell-eligible students apply for and receive the grant if the metric is included in the rating systems. Alternative metrics (e.g. income ranges rather than financial aid status) would require additional data collection by institutions, such as requiring all students to complete the FAFSA, although there are benefits to doing so for both students and institutions. We also recommend that the Department consider collecting and reporting additional data on the recipients of subsidized Stafford loans who do not receive Pell Grants (or another measure of lower-middle income students), and those who receive neither Pell Grants nor subsidized loans, as an additional nuance to the income distribution on campus. These data could be reported by institutions in IPEDS or populated in IPEDS for institutions to review using NSLDS.

² U.S. Department of Education (2012). *Fiscal Year 2013 Budget Request*. Retrieved from: <http://www2.ed.gov/about/overview/budget/budget13/justifications/p-sfa.pdf>

Progression and Completion: Who succeeds at an institution?

While information on the diversity of the student body helps provide context about an institution, data on student outcomes are just as important. When making a college decision, students need to know their chances of success at a particular institution to ensure their investment of time, money, and effort produces value in terms of a credential or productive transfer. At the same time, policymakers require data on what proportion of students complete, transfer, or are still enrolled – along with the total number of credentials awarded – to inform decisions about resource allocation and policy design. Institutions, in their continual efforts to improve, cannot do so without knowing their current level of performance with the students they serve.

It is critical that these outcome data be disaggregated by demographic characteristics, including *at least* race/ethnicity, income or financial aid receipt status, and enrollment (e.g. first-time, transfer) and attendance (e.g. part-time, full-time) status at entry. Too often, low-income students and students of color face lower chances of college success than their classmates.³ Comprehensive, accessible data can shine vital light on these gaps and drive action in a number of ways. Knowledge of low performance or disparities in performance can compel:

- Institutions to change policies and remove roadblocks to success.
- Policymakers to reward top performers or publicly shame poor performers.
- Students to “vote with their feet” by choosing colleges that are most likely to serve them well.

Data Availability for Reporting: Progression & Completion

While a base of data are available on student success through IPEDS’ retention and graduation rates, several important pieces of progression and completion information are missing from current collections. Table 2 includes a number of progression and completion measures, their current availability, and where improvements in data collection could be made. For one, collecting progression measures can provide crucial insights into student momentum, offering earlier indicators of student success than backward-looking completion measures. Currently, IPEDS includes first-year retention rates only, with no data on other progress measures like remedial completion rates, Satisfactory Academic Progress rates, and credits or time to credential. We recommend these data points be added to IPEDS to signal their importance in measuring student progress and to provide useful information to students and policymakers about how long it takes to attain a credential and what roadblocks (e.g. developmental education, academic progress) may stand in the way.

While only minimal data are available on progression, quite a bit of information is available on completion. However, these existing data suffer from a series of oft-lamented – yet highly fixable – limitations. Graduation rates offer the most notable example of a necessary data point in need of

³ Lynch, M. and Engle, J., “Big Gaps, Small Gaps: Some Colleges and Universities Do Better Than Others in Graduating Hispanic Students” and “Big Gaps, Small Gaps: Some Colleges and Universities Do Better Than Others in Graduating African-American Students.” Washington, D.C.: The Education Trust, August 2010.

improvement. The IPEDS graduation rate includes only first-time, full-time degree/certificate-seeking undergraduates and reports the proportion of those students who graduate within 100%, 150%, and 200% of time. These calculations omit part-time and transfer students, fail to account for upward transfers from community colleges as “successes,” and do not distinguish between associate and certificate completers.

Overall, about half (47 percent) of entering students are captured by the current first-time, full-time Graduation Rate Survey (GRS), but only 7 percent of institutions nationwide have a GRS cohort that includes less than 25 percent of entering students. In other words, only a small proportion of institutions have an entering class that is *grossly* underrepresented by the GRS cohort. Some sectors have far greater GRS coverage than others because students tend to enter these institutions as first-time, full-time students. For example, in fall 2011, 60 percent of students entering public and private nonprofit four-year institutions were captured in IPEDS graduation rates, while only about one-third of students at public community colleges and four-year for-profit institutions were included (33 percent and 30 percent, respectively).

The Department of Education has proposed new Outcome Measures, scheduled to be implemented in the 2015-16 collection cycle, that attempt to capture outcomes information on part-time and transfer students. While these new data will provide statistics on the outcomes of more students, they will not align or be comparable with the existing graduation-rate data because they differ in terms of completion timeframe and disaggregation of students by credential sought, race/ethnicity, and gender. They also will not produce the data necessary to sufficiently answer basic student progress and completion questions like:

- How many students transfer from a community college to a four-year institution?
- What is the graduation rate of bachelor’s-seeking students who enter an institution as transfers?
- What proportion of first-time, part-time community college students graduate within two, three, four, or five years?

Much of the higher education community believes “part-time and transfer student graduation rates are coming.”⁴ However, we expect there to be considerable dissatisfaction when the limitations of these new data are more widely known. We recommend that IPEDS be amended to align the new Outcome Measures with the existing Graduation Rate Survey (GRS) and retention rates. Not only will this alignment result in more usable data, but it also will reduce burden on institutions by streamlining the collection and reporting process through parallel definitions and methodologies (e.g. cohort development and tracking).⁵

⁴ In preparation for these comments as well as several projects on higher education data, IHEP has convened a series of meetings with experts in the higher education community around these very topics. In many cases, their “conventional wisdom” is cited throughout these comments.

⁵ For more on the limitations of the new Outcome Measures, see IHEP’s November 14, 2013 comments on the Outcome Measures: http://www.ihep.org/assets/files/radd/comment_on_ipeds_outcome_measures.pdf

In addition to the need for completion data that captures more students, policymakers and institutions also need better information on the success of students from different socioeconomic backgrounds. Under current law, institutions are required to disclose the graduation rates of Pell Grant recipients, subsidized Stafford loan recipients who do not receive Pell Grants, and students who receive neither Pell Grants nor subsidized Stafford loans.⁶ However, these data are not reported to IPEDS, so they cannot be evaluated on a broad basis for all institutions, and research has indicated that only a quarter of sampled institutions complied with the disclosure requirement when asked.⁷ Because institutions already are required to collect and disclose this information, we recommend incorporating it into the Graduate Rate Survey (GRS) as a disaggregate in the same manner as the race/ethnicity and gender disaggregates. While NSLDS is potentially an option for collecting and reporting graduation rates for Title IV recipients, it is not possible to compare those rates with non-recipients using that dataset, limiting its utility as a result.

⁶ "Information Required to Be Disclosed Under the Higher Education Act of 1965: Suggestions for Dissemination," National Postsecondary Education Cooperative, October 28, 2009, <http://nces.ed.gov/pubs2010/2010831rev.pdf>

⁷ Carey, Kevin and Kelly, Andrew P., "The Truth Behind Higher Education Disclosure Laws," Washington, D.C.: Education Sector, 2011, page 4.

Table 2. Progression and Completion: Who Succeeds at an Institution? Measures and Data Availability	
<i>Progression: Disaggregated at least by race/ethnicity, income or financial aid receipt, enrollment status (e.g. first-time, transfer), and attendance pattern (e.g., full-time, part-time) but ideally by all access measures above.</i>	
Measure/Data Element	Data Availability: Are these measures currently collected in IPEDS?
Remedial completion rate	Not currently available in IPEDS
Retention rate	Partially Available: IPEDS first-year retention rates are available disaggregated by full- and part-time status only.
Satisfactory Academic Progress Rate	Not currently available in IPEDS or any federal data collection
Credits to credential	Not currently available in IPEDS
Time to credential	Partially Available: Time to credential could be calculated for federal aid recipients using NSLDS with limitations (e.g. if students did not receive aid in their first term).
<i>Completion: Disaggregated at least by race/ethnicity, income or financial aid receipt, and enrollment status (e.g. first-time, transfer), but ideally by all access measures above.</i>	
Measure/Data Element:	Data Availability: Are these measures currently collected?
Completion/Graduation Rate	Partially Available: Only available for first-time, full-time students. For "other degree/certificate-seeking students" (non-bachelor's-seeking), IPEDS collects completion of a program of less than two years or a program of two but less than four years. These data do not clearly indicate whether the student completed an associate degree or certificate, nor do they align with the degrees conferred data in the Completions survey. New IPEDS outcome measures in 2015-16 will not make any distinction between credential levels (bachelor's, associate, and certificate).
Transfer rate	Partially Available: Currently, institutions with a transfer mission are required to report transfer-out data, and other institutions can report it voluntarily. The new IPEDS Outcome Measures will include a transfer-out rate for all institutions, but that transfer-out rate will not align with the GRS graduation rate and will not be disaggregated by level of receiving institution (upward vs. lateral vs. downward transfer).
Continued enrollment rate	Partially Available: IPEDS includes a still enrolled after 150% rate for less-than-two-year institutions and a still enrolled after 200% of time rate for four-year and two-year institutions. The new IPEDS Outcome Measures will measure whether students are still enrolled after 6 and 8 years, but will not align with the GRS outcomes as noted.
Degrees and certificates awarded	Currently Available in IPEDS

Potential PIRS Metrics: Progression & Completion

As with the access measures, the progress and completion measures listed above provide critical information to both students and policymakers about how well students – and which students in particular – are being served by which postsecondary institutions. As such, we recommend the data on those measures be improved as noted and made available to the public. However, for purposes of developing the rating systems, it is important to identify a few clear and compelling metrics for ease of presentation of the data in order to effectively impact consumer, institutional, and policymaker actions, the stated purpose of the systems.

Overall (150%) Graduation Rate

For degree programs, **overall 150% graduation rates** are currently collected and published, and are an easy to understand federal measure that does not require additional reporting requirements. The six-year window (for four-year institutions) and three-year window (for two-year institutions) provides flexibility for institutions while taking into account that many students take longer than 100% of time to complete. For consumer purposes, we actually recommend that the Department consider using 100% rates, since these resonate most with students' and parents' expectations per our understanding of findings from focus groups conducted by the Department on the Scorecard and other consumer information efforts. For accountability purposes, we recognize that 150% rates have achieved “consensus status” in the field to some extent given their ubiquitous use by institutions and policymakers. However, we caution against extending to 200% rates for purposes of the rating systems so as not to further dilute the strength of this metric.

As mentioned, we recognize that graduation rates only count full-time, first-time students who enroll in the fall, which for some schools account for a small proportion of incoming students. The graduation rate collection also does not factor in outward-transfer of students, or those who transfer into an institution and graduate. While new Outcomes Measures will include transfer-out as an outcome measure, the new measures will group students together regardless of the type of credential they are seeking (bachelor's, associate, or certificate), will not disaggregate transfer by level of receiving institution, and will not report a 150% completion timeframe for community colleges. As such, we do not recommend using the new Outcomes Measures in the rating systems for either consumer or accountability purposes.

While graduation rates are sometimes criticized as being heavily influenced by inputs – the academic strength of the incoming cohort – rather than the value-add of the institutions, research has indicated wide variety in graduation rates among institutions serving very similar populations.⁸ While the improvements in graduation rates noted above would undoubtedly strengthen their use in the rating systems, existing overall graduation rates remain a strong enough *signal* about whether the students that the research suggests have the best chances of succeeding at an institution – first-time, full-time students – do in fact complete their studies in a timely way.

⁸ Yeado, J. *Intentionally Successful: Improving Minority Student College Graduation Rates* (2013). Washington, DC: The Education Trust. Retrieved from: http://www.edtrust.org/dc/publication/intentionally_successful

Further, while we advocate for the improvements to graduation rates noted above, including tracking additional outcomes (e.g. transfer out and completion) as well as additional disaggregation (e.g. part-time, transfer), we do caution against the use of “outcome” or “success” measures without careful presentation. For instance, while it is certainly relevant and informative for students to know how many complete, transfer (and complete elsewhere), and remain enrolled displayed separately, we do not feel that it is appropriate to provide that information to students as a composite “success” rate. Doing so would be highly misleading, especially when comparing institutions with similar “success” rates but much different completion, transfer, or still enrolled rates. For the same reasons, it would be inappropriate to use a composite “success” rate for accountability purposes. Additionally, a transfer student’s eventual success at another institution certainly does reflect on the student, but not necessarily on the institution from which the student transferred (especially if the institution does not have transfer preparation as a major part of its mission), and should therefore be used cautiously as a measure of institutional success in either rating system.

Pell Grant Graduation Rates

While each of the disaggregates noted above provide important information to both students and policymakers, we recommend focusing on **Graduation Rates by Pell Grant Recipient Status** for purposes of the rating systems. The success of Pell Grant recipients is critical to promote social mobility, increased national attainment, and a strong return on the federal Pell investment. If added to the IPEDS collection, Pell graduation rates would provide a clear, well-targeted measure of how well institutions are graduating low-income students. It also would incentivize institutions to graduate populations that, currently, graduate at lower rates, thus focusing resources on those populations that stand to improve the most. Finally, this measure could be used as a protection to ensure that the best outcomes at an institution are not solely reserved for higher-income students, but that the students for whom an increasingly expensive college education is the biggest risk, namely low-income students, benefit as well.

As mentioned, Pell Grant graduation rates lag behind those of non-Pell recipients.⁹ Rather than Pell Grant graduation rates, institutions could be measured on the gap between the overall graduation rate and that of Pell recipients. However, measuring gaps can insert substantial complications into performance measures. Institutions can shrink gaps in a number of ways – some of which, such as decreasing non-Pell graduation rates, are undesirable. These issues can be accounted for in the rating systems; however, the simplest way to incorporate a measure of equity into the system is to measure and report the graduation rates of Pell Grant students separately.

⁹ For example, at four-year colleges, 60 percent of non-Pell recipients graduate within six years, compared with 43 percent of Pell recipients. Based on Institute for Higher Education calculations using the Beginning Postsecondary Students Survey (BPS 04:09). Calculations are based on beginning students who attend full-time in fall 2003. Pell status is based on fall 2003 Pell receipt.

Cost and Debt: How much do students invest in college?

At a time when college tuition and fees are increasing faster than inflation, family income, and healthcare costs,¹⁰ data on college costs are critically important to informing student decisions, public investment, and institutional policies. To make informed decisions, students need clear information about how much college will cost – not just their first year, but throughout their time at an institution – and how much they should expect to borrow. As college affordability continues to be an issue of great national concern, the rating systems should include measures of how much students (and their families) actually are paying and borrowing to attend college (See Table 3).

Data Availability for Reporting: Cost & Debt

One of the most pressing questions facing students and their families is how they will pay for college. As a result, college affordability and financial aid policy is one of the most salient challenges facing policymakers. Existing data provide a useful picture of the tuition and fees, cost of attendance, and net price that students will face their freshmen year, but they are left guessing about how much they will pay in subsequent years and about how much debt they should expect to accumulate throughout their college career. We recommend that IPEDS be amended to include cost information – tuition and fees, cost of attendance, and net price – not just for freshmen, but also for continuing students and transfer students. After all, college is at least a two or four-year investment for the vast majority of students, so families need access to more than one year of data. Of equal importance, cost for subsequent years is crucial for policymakers, especially if federal or state policy is to tie strings to measures of college cost. If sticks or carrots are associated with existing cost measures – those that only account for first-year costs – institutions could be incented to keep costs low for the first year, but raise them substantially in subsequent years, negatively impacting students.

The Higher Education Opportunity Act of 2008 required that institutions report net price data to IPEDS. These data, which represent what students pay for college *after* grant and scholarship awards, can help contextualize sticker price and provide a more realistic estimate of what they might pay. However, the existing net price data could be improved in several ways.

1. The average net price data are available only for students who receive grants or scholarships, downwardly biasing the results by omitting students who are paying full sticker price. The data underlying the average net price calculation do allow for calculating net price for all students *or* specifically for grant recipients. However, the most readily available net price data point is the average for only grant recipients, creating the potential for data-users, who may not recognize this nuance, to misinterpret the results. We recommend that NCEC calculate and report two net price figures using existing data – one for only grant/scholarship recipients and one for students regardless of aid receipt. This change would not require *any* additional reporting by institutions.

¹⁰ “College Costs Rising Four Times Faster Than Income, Two and a Half Times Faster Than Pell.” Washington, D.C.: Center for Law and Social Policy. <http://www.clasp.org/issues/postsecondary/pages/college-costs-rising-four-times-faster-than-income-two-and-a-half-times-faster-than-pell#sthash.EZrEvnlf.dpuf>

2. The net price data by income only include students who received Title IV financial aid, producing results that are fairly representative for low-income students, but far less representative more moderate and high-income students who are less likely to receive Title IV aid. It certainly is more difficult to obtain income information for students who do not receive Title IV aid, but several options exist for gathering these data. Some students fill out the FAFSA, but do not end up receiving Title IV aid, so institutions should be able to incorporate these non-Title IV recipients fairly seamlessly. For students who do not complete the FAFSA, institutions can survey students and families to collect income information. Survey data may be imperfect, but likely is sufficient to fill in data gaps. As previously noted, it is beneficial to both students and institutions to encourage more students to fill out the FAFSA anyway. Finally, to capture students for whom institutions simply cannot obtain income information, we recommend that IPEDS add an “income unknown” category to the net price data, ensuring all students – regardless of income data availability – are captured.

In addition to information on multi-year costs and more complete net price data, families and policymakers require far better information on student debt. The College Scorecard reports median borrowing using data from the National Student Loan Data System (NSLDS). However, these cumulative debt figures include completers and non-completers, producing sometimes confounding results. We recommend that institutions be required to report cumulative debt data to IPEDS, disaggregated by completion status, income/financial aid status, and race/ethnicity. The demographic disaggregates are crucial to informing students how much they likely will borrow and to informing policymakers and institutions about which students are being burdened most heavily with debt. For example, aggregate data show that African-American bachelor's degree recipients are far more likely than white, Hispanic, or Asian graduates to accumulate large debt loads (more than \$30,500) to pay for their education.¹¹ Clear data at the institution level can help clarify and spotlight these inequities, taking the first step towards closing them. To ease reporting burden on institutions, NCES eventually may be able to derive these data from NSLDS using institutionally reported completion information, on which NCES recently issued additional guidance.¹² Once this new completion information has been tested and validated, it can be used to calculate cumulative debt data for completers, non-completers, and students of different financial aid statuses. However, NSLDS does not include information on race/ethnicity, so to paint a complete picture of student borrowing, institutions will need to continue reporting debt by race/ethnicity to IPEDS or add that element to NSLDS.

¹¹ According to NPSAS data from 2007-2008, 27 percent of Black bachelor's degree recipients borrowed \$30,500 or more, compared with 16 percent of white recipients, 14 percent of Hispanic recipients, and 9 percent of Asian recipients. (Baum, S. and Steele, P., “Who Borrows Most? Bachelor's Degree Recipients with High Levels of Student Debt.” New York, NY: The College Board, April 2010).

¹² In 2012, NCES issued “important enrollment reporting reminders,” noting “We have been especially concerned that some schools are not complying” with the requirement to report completion/graduation data. The guidance reminds schools of codes to use for students' enrollment statuses, e.g. “graduated” and “withdrawn.” (“NSLDS Enrollment Reporting Process, Attachment to GEN-12-06,” March 30, 2012, <http://www.ifap.ed.gov/dpclatters/GEN1206.html>).

Table 3. Cost and Debt: How Much Do Students Invest in College? Measures and Data Availability	
Measure/Data Element	Data Availability: Are these measures currently collected in IPEDS?
Tuition and Fees	Available: Tuition and fee data are reported in the IPEDS Institutional Characteristics (IC) survey. In-state and out-of-state tuition and fees are reported for first-time, full-time undergraduates. Average tuition and fees are reported for all undergraduates. Tuition and fee data are not disaggregated for transfer or continuing students.
Cost of Attendance	Partially Available: In-state, in-district, and out-of-state cost of attendance are reported only for first-time, full-time degree/certificate-seeking undergraduates by living status (e.g. on campus, off-campus with family, off-campus not with family) in the IPEDS Institutional Characteristics study.
Net Price	Partially Available: Average net price data are available for first-time, full-time undergraduates who receive grant or scholarship aid. Net price data are disaggregated by income bands for first-time, full-time undergraduates who receive Title IV aid. Both of these net price data points omit students paying out-of-state tuition (at publics), transfer and continuing students, and students who do not receive financial aid (either Title IV or grants/scholarships).
Cumulative debt, disaggregated by race/ethnicity, income or financial aid receipt, and completion status	Partially Available: The College Scorecard reports total federal loan debt (including Parent PLUS loans) among students leaving an institution, using NSLDS. It does not separate completers from non-completers, disaggregate by type of federal loan debt (e.g. PLUS), or include private loan debt.

Potential PIRS Metrics: Cost & Debt

Again, each of the cost measures listed above provide critical information to both students and policymakers about how much students must invest in their college education. As such, we recommend the data on those measures be improved as noted and made available to the public. However, for purposes of developing the rating systems, it is important to identify a few clear and compelling metrics for ease of presentation of the data in order to effectively impact consumer, institutional, and policymaker actions, the stated purpose of the systems.

Average Net Price

An average net price figure – improved by including non-grant and scholarship recipients – would at least provide an estimate of what students actually are liable for when paying college expenses. Some definitional changes would be necessary if strings are to be attached to this measure, such as also including transfer and continuing students. However, even as currently structured it would provide a better picture than sticker price figures – including cost of attendance – in the rating systems. Net price is directly influenced by institutional and state policies, through both tuition setting and provision of grant aid. By definition, net price gives credit to institutions for their grant policies by focusing on what students actually pay instead of what is charged (or “sticker price”).

Net Price for Low-Income Students

Additionally, we recommend using a metric that measures institutions on the net price for students from families making under \$30,000, which is currently available in IPEDS. Like average net price, this metric would focus on what students actually pay after grant aid. Unlike average net price, this measure would give credit to institutions who keep net prices affordable for low-income students. As with other IPEDS data, this data point only measures price for first-time, full-time students, which could induce institutions to keep costs low for the first year, but raise them substantially in subsequent years unless this nuance is addressed. Ideally, the data would be collected and a net price metric would be disaggregated by enrollment (e.g. first-time, transfer, continuing) status in order to prevent institutions from making tuition increases in later years, but at minimum continuing students should be included in the calculation.

Additionally, it currently is not possible to separate the aid distributed by institutions, states, and the federal government in this metric, meaning institutions could alternately be given credit for state (e.g. low tuition) or federal (e.g. increasing Pell Grants) policies, or be harmed when states or the federal government pull back from support for low-income students. However, these interactions between policies at various levels are not necessarily problematic. Holding a federal lever over institutions for net price – which is influenced by state policy – could help nudge states toward stronger investments in higher education and need-based grant aid and give institutions more room to pressure their state to maintain investment. And, the benchmarks within the rating systems could be set in a way that changes in federal aid do not trigger federal penalties for institutions. For example, if Pell Grants were reduced by \$100, the net price threshold could be reduced by \$100 as well.

Cumulative Debt

Finally, we recommend including a cumulative debt measure that shows the average amount of debt a student takes on to attend an institution. A cumulative debt measure would serve a similar purpose to a net price measure, but would more explicitly separate how much debt students take on at an institution from other payment methods. Like net price metrics, cumulative debt is directly influenced by state and institutional policies on tuition and grant aid.

Currently, the College Scorecard includes a cumulative debt figure that includes all federal loans – including Parent PLUS borrowing. It does not separate completers from non-completers, disaggregate by type of federal loan debt, or include private loan debt. We recommend disaggregating by type of loan and completion status for a few reasons. First, loan terms vary in generosity, and it would be worth examining which institutions require students to take on more onerous forms of debt (such as Parent PLUS loans and private loans, neither of which are eligible for income-based repayment protections). We do recognize, however, that there are significant limitations for institutions to obtain complete information on private loans.

Second, by not disaggregating by completers and non-completers, the data can produce confounding results. An institution with high dropout rates and high costs can show a similar median debt number as a low-cost institution with high graduation rates simply because students are enrolled in the high-cost institution for a shorter period of time. To prevent the use of confusing information, we recommend



that institutions be required to report cumulative debt data to IPEDS, disaggregated by completion status, income/financial aid status, and race/ethnicity.

Repayment and Earnings: How do students fare after college?

Measuring the personal economic return on investment of college requires data on several components: completion and cost – as discussed above – along with success after college. In other words, “value” measures what a student and society receives in return for the public and private investment in higher education. While, on the whole, the evidence is clear that college produces a sound return on investment,¹³ outcomes vary by institution and program.¹⁴ It is impractical to expect students and their families to make informed financial decisions about how much to pay and how much to borrow for college without some basic information on post-college employment, short- and long-term earnings, and loan repayment. Similarly, these types of data can help policymakers identify institutions that are preparing their students well for the workforce and for loan repayment and those that are setting their students up for financial difficulties. By examining workforce and loan repayment data, institutions can tweak program content, adjust job counseling policies, or rethink connections with industry that could benefit their students in the long-run.

Equally important to measuring college value is how much students learn at a particular institution. While graduation rates are a basic measure of quality, and employment outcomes attempt to approximate what students gain in the labor market from a program of study, learning outcomes may offer a more true measure of college quality. The purpose of college, after all, is to impart learning upon students to help them to succeed in the workforce and society. A variety of efforts and tools, including the Degree Qualifications Profile, the Voluntary System of Accountability, and the Collegiate Learning Assessment are attempting to measure learning in practical ways. This area is ripe for research, and as these efforts continue to develop, measures of student learning should be incorporated into publicly available datasets to inform students, policymakers, and institutions, and potentially, future iterations of a ratings system.

Data Availability for Reporting: Repayment & Earnings

Perhaps the area in which postsecondary information is most severely lacking is post-college outcomes. While we certainly need better data on college access, completion, and affordability, our current data systems provide at least some directional data on these topics. However, in the case of how students fare after college, students, policymakers, and institutions currently have broad access to only one institution-level data point: cohort default rates (CDRs). While CDRs do provide useful information about how many students are facing severe difficulty in repaying their loans, they do not provide any indication of the success of non-defaulted students, some of whom may be avoiding default, but still

¹³ Carnevale, A., Rose, S. and Cheah, B., “The College Payoff: Education, Occupations, Lifetime Earnings.” Washington, D.C.: The Georgetown University Center on Education and the Workforce, August 2011; Baum, S., Ma, J. and Payea, K., “Education Pays 2013: The Benefits of Higher Education for Individuals and Society.” New York, NY: The College Board, October 2013.

¹⁴ “2011 Gainful Employment Informational Metrics,” Federal Student Aid, U.S. Department of Education, <http://studentaid.ed.gov/about/data-center/school/ge/data>.

struggling financially. Such borrowers include those who are delinquent without defaulting – about a quarter of all borrowers.¹⁵

A variety of measures on student outcomes after college can provide a better understanding of institutional value-add. For instance, information on employment, earnings, and loan repayment can indicate how successful students are *after* attending an institution. In our current system, however, students cannot compare the expected earnings returns of different institutions or programs to identify which colleges will give them the best chance at a ticket to the middle class, especially in relation to expected debt levels.

Several options exist for collecting and reporting workforce data, including five major sources: Social Security Administration (SSA); Unemployment Insurance (UI) records, which could be used in conjunction with the Federal Employment Data Exchange System (FEDES); National Directory of New Hires; and the Longitudinal Employer-Household Dynamics (LEHD) program. Each source has different strengths and weaknesses, as outlined in Table 4. For example, the SSA data include federal employees and the self-employed – two groups omitted from UI records – but are only available annually, while UI data are available quarterly. If a state-based approach is implemented, it is important that states continue to improve cross-state data sharing and that federal rules and guidance permit and encourage such sharing, through programs like WRIS2. Under a state-based model, this sharing will be critical to ensure that programs whose graduates typically find work in another state – either within the region or otherwise – have their outcomes fairly represented in the data.

Though none of these data sources are fully exhaustive and each presents its own complexities and limitations, the most direct approach to acquiring workforce data at the institution or program level is to link SSA earnings information with student-level data submitted to the Department of Education by all postsecondary institutions. The SSA already has participated in data matches with the National Student Loan Data System (NSLDS) – which only includes federal financial aid recipients – to generate earnings information as part of Gainful Employment, indicating that such matches are technically feasible. A move toward a comprehensive student unit record system could provide more comprehensive, high-quality data on workforce results.

¹⁵ Cunningham, Alisa and Kienzl, Gregory (2011). *Delinquency: The Untold Story of Student Borrowing*. Washington DC: Institute for Higher Education Policy. Retrieved from: http://www.ihep.org/assets/files/publications/a-f/Delinquency-The_Untold_Story_FINAL_March_2011.pdf

**Table 4. Data Sources for Employment and Earnings Results
Advantages and Disadvantages**

Data Source	Advantages	Disadvantages
UI Wage Data	<ul style="list-style-type: none"> • Includes all workers covered by the Federal Unemployment Tax Act (FUTA); approximately 89 percent of the civilian labor force. • Data are quarterly, which provides more flexibility in creating employment and earnings metrics. • Generally includes industry, which allows a metric for employment in an industry related to the field of study. 	<ul style="list-style-type: none"> • Does not include self-employed, military, federal civilian, postal employees, railroad employees, and a few others. • In most states does not include start date, hours worked, or occupation. • Data are “owned” by states, and therefore requires additional data exchange process to obtain data from multiple states (WRIS, WRIS2). • States vary substantially in access policies, and some have been very restrictive.
National Directory of New Hire data	<ul style="list-style-type: none"> • Includes all workers covered by UI, plus military and federal civilian employees. • Data are quarterly, which provides more flexibility in creating employment and earnings metrics. • Includes UI claimant information (unemployed). 	<ul style="list-style-type: none"> • Does not include self-employed. • Does not include hours worked, industry or occupation. • Not currently permitted to be used to support calculation of outcomes for postsecondary institutions.
Social Security Administration (SSA) earnings data	<ul style="list-style-type: none"> • Includes essentially all workers: those covered by UI, and those exempt from UI, including federal civilian, military, and self-employed. • Data are obtained from IRS and maintained centrally by SSA. 	<ul style="list-style-type: none"> • Data are annual only (for Master Earnings File). • Currently matches are restricted to data for students submitted through the National Student Loan Data System (NSLDS). This is not an inherent limitation of the SSA data, but is a limitation of the availability of student data with SSNs. • Currently does not include industry or occupational codes.
Federal Employment Data Exchange System (FEDES)	<ul style="list-style-type: none"> • Includes military and federal civilian employees.¹⁶ 	<ul style="list-style-type: none"> • Does not include self-employed. • Includes 37 states and DC.¹⁷ • Use of FEDES is restricted.¹⁸

¹⁶ United States Postal Service (USPS) employees were originally included but are not currently; negotiations are underway to resume matching for these workers.

¹⁷ The states participating in the most recent match were: Alabama, Alaska, Arkansas, Arizona, California, Colorado, Connecticut, District of Columbia, Florida, Georgia, Hawaii, Idaho, Illinois, Indiana, Kansas, Kentucky, Maine, Maryland, Minnesota, Missouri, Montana, New Jersey, New Mexico, New York, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, South Carolina, South Dakota, Tennessee, Texas, Vermont, Virginia, Washington, West Virginia, and Wisconsin.

¹⁸ Currently use of FEDES is restricted to “Satisfying, or contributing to, Federal performance measurement and consumer report activities required by the United States Office of Management and Budget (OMB) or federal law or regulation, or satisfying, or contributing to, State performance measurement and reporting requirements authorized under state law or regulation.”

Table 4 (continued). Data Sources for Employment and Earnings Results		
Advantages and Disadvantages		
Longitudinal Employer-Household Dynamics (LEHD)	<ul style="list-style-type: none"> • Includes all workers covered by UI, plus military and federal civilian employees. • Quarterly data. • Provides useful information on the local labor market context, including employment and earnings by industry and by education, gender, ethnicity, age, firm size and firm age. 	<ul style="list-style-type: none"> • Due to access restrictions, this is not a source of data on results for specific sets of students. • Does not include self-employed.
<p><i>Source: Analysis by the Center for Law and Social Policy (CLASP) as a part of the Reimagining Aid Design and Delivery Project, 2014 (forthcoming)</i></p>		

In addition to basic data on post-college employment and earnings, students and policymakers need to know whether students are able to repay their loans without undue burden. Measures such as repayment rates or repayment progress ratios can provide a sense of how successful students are at repaying their debts. Repayment rates measure the percent of students, dollars, or institutional loan portfolio that are “in repayment,” defined as having a balance that is decreasing over time. In a similar but more nuanced way, a repayment progress ratio measures the proportion of students on track to repay their loan in a set amount of time (e.g. 10 years) to indicate whether students are earning sufficient income to make substantial progress in repaying their debt. Loan repayment measures, in conjunction with labor market outcome data can provide students, families, policymakers, and institutions with crucial information to inform decision-making and policy development.

Table 5 outlines the current availability of various workforce data and Table 6 outlines repayment metrics, as well as data availability. For the most part, repayment measures have been or could be collected through NSLDS, though could be improved if disaggregated by completion status, socioeconomic status, race/ethnicity, or program of study.

Program-level data may indeed be more important when measuring employment, earnings, and repayment than in other categories (such as completion). On one hand, holding programs accountable for their performance would offer a more fine-grained and targeted system of checks and balances, assuring that institutions offer programs with real value in the labor market. However, implementing a program-level accountability system would require far more in-depth data collection than currently exists. Program-level data on debt and repayment were collected under Gainful Employment, but expanding these types of accountability measures to all institutions would require a substantially more robust data system, such as a student unit record system.

Also, tying accountability to program-level debt or repayment measures may encourage the use of differential tuition, in which majors with high expected earnings would charge more. Extreme program-level tuition differentiated could stratify program access based on family income levels, making it more

difficult for low-income students to access high paying fields. Finally, many program-level measures likely would suffer from small sample sizes, making it difficult to measure performance with confidence. At minimum, institution-level measures allow colleges flexibility in meeting the benchmarks. If the Department uses institution-level data, the performance of students in programs that tend to produce high wages, for instance, can balance out the performance of students in programs that tend to produce low wages. Ideally, the institution will ensure that *all* of their students are successful – in college and beyond – but the institution-level accountability measures will allow them to offer programs in a variety of fields with a variety of expected earnings and outcomes.

Table 5. Earnings and Workforce Outcomes: How Do Students Fare After College? Measures and Data Availability	
Measure/Data Element	Data Availability: Are these measures currently collected?
Workforce success measures: Disaggregated by completion status, socioeconomic status (while in college), race/ethnicity, and program.	
Employment Rate (minimally), but could consider Employment Retention Rate, Related Employment Rate as well	Partially Available: Employment data could be collected from Social Security Administration (SSA), Unemployment Insurance (UI) wage records in conjunction with the Federal Employment Data Exchange System (FEDES), or the National Directory of New Hires (NDNH). Data in one of these other federal systems would need to be linked to student-level information from institutions of higher education, which is only available through NSLDS or the National Student Clearinghouse (NSC).
Initial Median Earnings (minimally), but could also consider Subsequent Median Earnings, Earnings Change over Time	Partially Available: SSA calculated the mean/median earnings of program graduates in the third and fourth year after completion as part of Gainful Employment in 2011. Earnings data also could be calculated from UI wage records, in conjunction with FEDES, or NDNH.

Table 6. Repayment: How Well Can Students Pay off Loans? Measures and Data Availability	
Repayment Measures: Disaggregated by completion status, socioeconomic status, race/ethnicity, program	
Cohort default rate	Partially Available: The Department of Education (ED) currently calculates cohort default rates using the National Student Loan Data System (NSLDS), but does not disaggregate the results by completion status, socioeconomic status, or race/ethnicity. Also does not include PLUS or private loans.
Repayment rate	Partially Available: ED has calculated repayment rates using NSLDS as part of gainful employment, but does not calculate and release them on a regular basis. ED has not disaggregated these rates by completion status, socioeconomic status, or race/ethnicity. Thus far, repayment rates have not included Parent PLUS or private loans.
Repayment progress ratio (% of students on schedule to repay their loans within 10 years)	Partially Available: Repayment progress ratio has not been calculated in the past, but could be calculated using NSLDS.
Debt-to-earnings ratio	Partially Available: ED has calculated debt-to-income ratios in the third and fourth year after completion through collaboration with SSA under Gainful Employment. These data were reported only for completers and were not disaggregated by socioeconomic status or race/ethnicity, and they are not currently scheduled to be reported on a regular basis.
Graduate school preparation/enrollment rate	Partially Available: Four-year institutions must disclose to students (but not report more broadly) "the types of graduate and professional education in which graduates...enrolled," ¹⁹ but not enrollment rates or graduate school entrance exam scores.

Potential PIRS Metrics: Repayment & Earnings

Repayment Rate

Repayment rates represent the ability of borrowers to make minimum progress against the principal of their loan balance. Repayment rate has been defined by the U.S. Department of Education, either in statute or proposed regulations in two ways. In initial rounds of Gainful Employment rules released in 2011, repayment rate was defined as the percent of dollars in “repayment” (in which principal was reduced by at least \$1),²⁰ with institutions or programs at risk of sanction if a certain percent of dollars are not in “repayment.” In other words, if a student pays at least \$1 toward principal, then their loan balance is considered in “repayment.” Under original Gainful Employment regulations, an institution passed the repayment rate metric if at least 35 percent of loan dollars were in repayment. However, in

¹⁹ HEOA Sec. 488(a)(1)(S)

²⁰ U.S. Department of Education (2011). Program Integrity: Gainful Employment-Debt Measures. 76 FR 34385. Retrieved from: <https://www.federalregister.gov/articles/2011/06/13/2011-13905/program-integrity-gainful-employment-debt-measures#h-26>

2012 a federal judge found that the Department did not provide sufficient justification for this 35 percent threshold, and the regulation was nullified.²¹

In November 2013, the Department of Education proposed for discussion a new definition of repayment rate, which measures whether or not an institution's total loan portfolio is in repayment or not. Programs "fail" this repayment measure if the principal of their entire loan portfolio does not decrease by at least \$1 over the course of the year.²² Under both definitions, repayment rates were to include program completers and non-completers.

There are a number of advantages to using repayment rates as a metric. Repayment rates reflect the ability of students to make more than interest-only payments on their loans after a given amount of time. Further, by including non-completers in the calculation, repayment rates are also a proxy for completion, since borrowers who earn a credential are more likely to be able to pay off loans. In fact, non-completion is consistently the best predictor of student loan default.²³ Finally, unlike CDRs, borrowers in deferment or forbearance count as negative outcomes in repayment rate calculations.²⁴ If a student enters deferment or forbearance, her loan balance will not decline, so her balance will be counted as not in "repayment." Because of this treatment of deferment and forbearance, institutions cannot use these tools to avoid repayment rate-based sanctions as they can with CDRs.

However, there are some limitations to using repayment rates in the rating systems – primarily related to their lack of comprehensiveness in identifying struggling borrowers. Without making a distinction between those paying back \$1 of principal and those making a sizeable dent in their debt, an institution's borrowers could be considered in "repayment" while not making substantial progress to pay off loan balances. Paying off \$1 of principal should be a bare minimum expectation, but truly successful borrowers should be able to make more progress.

In addition, the Department may want to consider disaggregating repayment rates by completion status to provide a more nuanced picture of how well an institution's graduates are doing in loan repayment. Also, repayment rate definitions have not traditionally accounted for Parent PLUS or Perkins loan borrowing (or private borrowing, for that matter). As a result, repayment rate measures do not hold institutions accountable for the full amount of debt students or families incur and their ability to repay

²¹ APSCU v. Arne Duncan and the U.S. Department of Education, 2012. U.S. District Court for the District of Columbia. Civil Action 11-1314 (RC), Retrieved from: https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2011cv1314-25

²² U.S. Department of Education (2013). Overview of Draft Gainful Employment Regulation Negotiated Rulemaking Session 2 (PowerPoint). Draft for discussion purposes 11.8.13. Retrieved from:

<http://www2.ed.gov/policy/highered/reg/hearulemaking/2012/gainfulemployment.html>; Ben Miller has proposed a similar repayment measure. See: Miller, Ben (2013). Improving Gainful Employment: Suggestions for Better Accountability. New America Foundation. Retrieved from: <http://www.newamerica.net/sites/newamerica.net/files/policydocs/Improving%20Gainful%20Employment%20FINAL.pdf>

²³ Jacob P. K. Gross, Osman Cekic, Don Hossler, and Nick Hillman (2009). What Matters in Student Loan Default: A Review of the Research Literature. *Journal of Student Financial Aid* 39:1. Pp. 19-29

²⁴ In the 2011 Gainful Employment rule, three percent of loan balances could count as "in repayment," even if they were negatively amortizing under an income-based plan. Such a carve-out reduces institutional accountability unnecessarily. For more on the relationship between income-based plans and repayment rates, see: Miller, Ben (2013). Do Income-Based Payment Plans Really Ruin Repayment Rates? New America Foundation. Retrieved from: <http://inthetank.newamerica.net/blog/2013/12/do-income-based-payment-plans-really-ruin-repayment-rates>

it. Perkins loans could and should be included in repayment measures, but data on repayment of private loans are not available, and it would require additional discussion regarding incorporating Parent PLUS loans into the repayment metric because a different individual (the parent) actually takes out these loans. We recommend that Perkins loans be added to repayment rates, though.

Repayment rates are not calculated regularly and reported by the U.S. Department of Education; however the data to do so do exist in NSLDS, and the Department has calculated and released repayment rates several times as part of the Gainful Employment negotiated rulemakings. These data releases have calculated various definitions of repayment rates to model the impact of proposed regulations. Because these rates can be calculated using NSLDS, using them in rating systems would require no increase in reporting burden for institutions.

Repayment Progress Ratio²⁵

Alternatively, a Repayment Progress Ratio provides another view on how much of a dent students are able to put in their loan balances upon leaving school. Rather than examining whether or not a borrower is paying down principal by \$1, a Repayment Progress Ratio would look at the proportion of loans “on-track” to be repaid over a certain period of time. “On-track” could be defined in various ways. For example, the Department could consider the percent of loans on-track to be repaid within 10 years, since around two-thirds of federal direct loan borrowers are currently enrolled in a 10-year repayment plan.²⁶

The major benefit of this type of measure is that it is more nuanced than current repayment rates proposed by the Department since it would measure whether students are making substantial progress toward paying down loans, rather than the minimal amount of progress used as a standard under repayment rates. This measure likely would produce similar results as the debt-to-earnings ratios, but would maintain a direct tie to loan repayment. It also could include non-graduates, which debt-to-earnings ratios do not currently do. The Department could calculate Repayment Progress Ratios with the same information as was used for repayment rate, via NSLDS. The major drawback of a measure of repayment progress is that it likely will be more complicated than a repayment rate.

Minimum Earnings Threshold

As mentioned, workforce outcomes data are not collected routinely in a way that would lead to easy incorporation into the rating systems. However, we recommend that the Department consider a minimum earnings guideline for accountability purposes to ensure that students are receiving a minimal level of value from an institution, as demonstrated by labor market returns. These metrics could include: percent of graduates earning more than minimum wage; percent of graduates who earn at least 200 percent of the poverty level, or percent of graduates who earn more than the amount that a high school graduate in the same field earns. For consumer information purposes, however, students should know

²⁵ The concept of a repayment progress ratio arose from discussions between IHEP and Mark Kantrowitz, Senior Vice President and Publisher of Edvisors Network.

²⁶ Rohit Chopra (August 5, 2013). “A Closer Look at the Trillion.” Washington DC: Consumer Financial Protection Bureau. <http://www.consumerfinance.gov/blog/a-closer-look-at-the-trillion/>

in substantially more detail what they can expect to earn after attending school (e.g. at one, five, and 10 years out), so they can compare their projected income against the investment they are about to make in an institution.

Workforce outcomes measures should also incorporate – and disaggregate by – completers and non-completers, since any investment by a student in an institution should result in a minimal level of value in the labor market.

Access	Percent Pell
Progression and Completion	Overall Graduation Rate Graduation Rate for Pell Grant Recipients (100% for Consumers, 150% for Accountability)
Cost and Debt	Average Net Price (including non-aided students) Average Net Price for Low-Income Students Cumulative Debt (disaggregated by completion)
Repayment and Earnings	Repayment Rate OR Repayment Progress Ratio Earnings (Minimum Threshold for Accountability; Averages for Consumers)

2.0 & 3.0 Weighting and Scoring and the Development of Comparison Groups

Decisions regarding weighting, scoring, and institutional comparisons will have a critical impact on the results of the rating systems. While these design elements are highly technical, they will produce the evaluative and contextualized results aimed at impacting students and policymakers. Because of the significant bearing that these decisions will have on the rating systems’ results, consumer behavior, and funding policy, we recommend that the Department carefully and thoroughly test and examine the consequences of various options.

While the aforementioned metrics (see Table 7) might be able to serve both purposes – consumer information and institutional accountability – decisions around weighting, scoring, and comparison must differ depending on the audience and purpose, thereby necessitating separate rating systems for these dual purposes in our opinion. Some technical options may make sense in an accountability context, but produce misleading or overly complex information for students. This section summarizes technical recommendations for the design of a consumer-centric rating system and separately for various types of institutional accountability systems.

Rating System for Students

For consumer purposes, we recommend that the rating system provide unadjusted information to students on key metrics using easy-to-understand terms in graphical formats that allow for side-by-side comparisons. As with the current College Scorecard, students should be able to compare the institution’s outcomes to all institutions using averages, ranges, and terciles, as appropriate. Students also should be able to align institutional results in a side-by-side format, encouraging and simplifying



comparison shopping, which they currently cannot do with either the College Scorecard or College Navigator. In designing a consumer-focused ratings system, we recommend that the Department:

- Present and rate each metric, but not a composite rating
- Keep institutional groupings broad and not adjust rating cutpoints on metrics
- Complement the ratings with information on similar, but better-performing institutions
- Invest in counseling: The best consumer information system

Present and rate individual measures, but not a composite rating

Each of the potential PIRS metrics discussed above and summarized in Table 7 are relevant for students and their families as they make college choices, and students need clear information on *each* of these indicators. It may be tempting to boil institutional performance on a variety of indicators down to a single number or rating, but we feel that doing so would be a mistake. A stand-alone composite measure, despite its seeming simplicity, would lack intrinsic meaning and mask the critical underlying data elements, thereby reducing its influence on consumer behavior. When making college choices, students need to balance a variety of factors – such as cost and outcomes – against each other, and make value-driven, personalized decisions based on that amalgam of information. A composite indicator undermines that decision-making process by making critical decisions *for* students about which indicators are most important, particularly if the indicators are weighted, and which may not align with the students’ interest or needs. For more details on our concerns about the technical pitfalls composite indicators, see the discussion of composite indicator options in the “Rating System for Accountability” section.

Even without a composite score, a consumer information tool could be designed in a way that allows students to compare schools on more than one factor at the same time. For example, the tool could generate a grid that visually represents institutional performance on two factors for multiple schools, allowing students and parents to make their own evaluations about the results. See Figure 1 for an example of this type of tool.

Figure 1: Visuals can help students evaluate multiple measures at the same time



At minimum, we recommend that the Department continue to rate institutions on each of the metrics, as it does on the current College Scorecard. Ratings, such as “low,” “medium,” and “high,” help contextualize numbers that otherwise may be difficult for students to interpret on their own (for example, how good is a 62 percent graduation rate?). Furthermore, each institution should be rated on each indicator in a way that allows students to compare across colleges of interest easily. Other college search tools, such as College Reality Check, College Results Online, and College InSight, allow for side-by-side comparisons, and both winners of the recent Lumina Foundation sponsored competition to redesign the College Scorecard included this comparative feature.²⁷ We recommend that a consumer-focused rating system make this basic enhancement to the Scorecard (see Figure 2 for an example).

²⁷ Jillian Klein of Capella University and Rachel Fishman and Ross Van der Linde of New America Foundation won Lumina’s *College Scorecard Design Challenge*. See: Innocentive and Lumina Foundation Announce Winners of College Scorecard Design

Figure 2. Consumer Tools should allow side-by-side comparisons



If the Department does pursue developing and publishing a composite rating for consumer purposes, we strongly recommend that (1) the individual metrics still be made available separately, (2) the individual metrics are rated separately with unadjusted cutpoints (see more below), (3) the individual metrics are rated in such a way that aligns with the composite rating (e.g. the same scale), (4) only the most critical metrics (e.g. cost and completion) are included in the composite so as to avoid “double-counting”

Challenge (2013). Press Release. Lumina Foundation. Retrieved from:
[http://www.luminafoundation.org/newsroom/news_releases/2013-12-19.html#!prettyPhoto\[pp_gal1\]/1/](http://www.luminafoundation.org/newsroom/news_releases/2013-12-19.html#!prettyPhoto[pp_gal1]/1/)

metrics that covary in the same direction, and (5) the weighting scheme is made highly transparent and possibly even adjustable by students through the interface.

Keep institutional groupings broad and do not adjust rating cutpoints on metrics

Currently, the Scorecard only distinguishes between institutions that primarily grant bachelor's degrees and those that primarily grant associate degrees. We feel that this limited distinction is entirely appropriate, as bachelor's and associate-granting institutions serve different purposes and the definitions underlying their completion metrics in particular are sufficiently different to warrant separate treatment. For institutions at which both bachelor's and associate degrees are prevalent, we recommend that *two* Scorecards be generated – one for each type of credential so students can evaluate the college based on the type of credential they intend to pursue.²⁸

However, we recommend that the rating system not use any further distinction beyond primary award granted to identify “similar colleges” in a consumer context. Using any more nuanced methodology for a consumer-centric rating system could create confusing and misleading results and constrain student choice sets, missing an opportunity to expand students' college searches, which is a stated goal of the current Administration.

Confusing and misleading results

We feel that creating different cutpoints for “low,” “medium,” and “high” performance for different sets of institutions based on their student or institutional characteristics would create confusing information at best, and misleading information at worst. Imagine the confusion a student may face if she reviews the Scorecard for two colleges and finds that one 70 percent graduation-rate institution is rated as “medium,” while another 40 percent graduation-rate institution is listed as “high.” Setting different cutpoints based on comparison groups that are any more limiting than the existing Scorecard grouping could generate this type of confusing information for students.

Perhaps more dangerous is the potential for inadvertent lowering of expectations for students, who may search ratings or Scorecard information for a variety of different types of colleges for a variety of different reasons. Consider a high-achieving student (or even a student with modest academic qualifications) who lacks hands-on counseling from an adult, but seeks information on her own. She searches for the one college she knows – a \$23,000 net price institution with a 22 percent graduation rate located in her hometown. This cost and outcomes information shows that the institution does not provide a good value for most students who attend. However, if that college is only compared with a limited set of narrowly-defined “similar” institutions, the high net price may be rated as “low” or “medium” and the low graduation rate may appear as “medium” or “high,” leading the student to believe the institution is likely to serve her better than the raw numbers indicate. But, providing an

²⁸ Four-year institutions report to IPEDS both graduation rates for bachelor's-seeking students and for other degree/certificate-seeking students. Completions for the other degree/certificate-seeking cohort include completers of programs of less than two years, completers of programs of two, but less than four years, and completers of bachelor's or equivalent degrees. As more institutions, especially community colleges and for-profit institutions, offer a broader range of credential types, the National Center for Education Statistics may need to offer further guidance on proper reporting of multiple cohorts.

honest and unvarnished evaluation of this institution’s performance (poor graduation rate and high cost) could encourage the student to investigate other colleges that may serve her better. A system with adjusted cutpoints could be particularly damaging to student expectations if it signaled lower standards for low-income or underrepresented students. Students should choose the *absolute* best college they can get into, (and afford), not the “best” college adjusted for “inputs.”²⁹

Constraining student choice sets

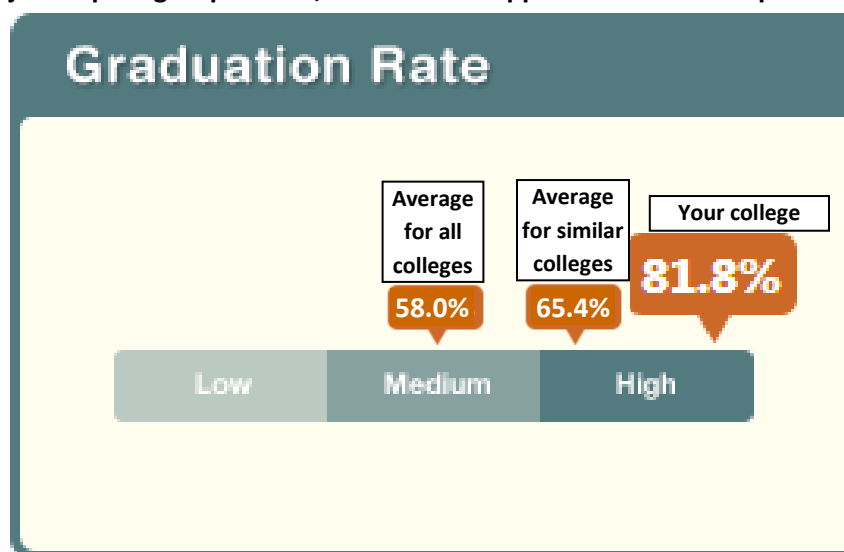
Students create college choice sets using an assortment of factors, and we feel that any type of federally defined institutional grouping should be based on what matters to *students, not institutions*. A single student may consider a mixture of institution types – publics/privates, research/liberal arts, PWIs/HBCUs, big/small, etc. Because of the diversity of college groupings used by students, designing a consumer rating system that creates comparison groups based on variables like sector, size, or student demographics could artificially limit student choice sets. For example, sector should be a completely irrelevant variable for students, and in some cases, students may not even know – or need to know – the difference between a public, private nonprofit, and private for-profit institution. They need to know how much college will cost and whether they are likely to graduate and be successful after college. Rather than accounting for irrelevant distinctions based on institutional or student characteristics, the rating system should prioritize the Department’s stated goal of promoting *value*, based on cost and student outcomes. To accomplish this, the rating system should be designed in a way that encourages students to consider a broad array of colleges, rather than a constrained choice set. In fact, one potential benefit of a consumer information system is to open students’ eyes to high-value institutions that they had not previously considered.

Because of the potential for confusing and misleading results and constrained choice sets, we feel that adjusted cutpoints or limited comparison groups are not appropriate for a rating system that provides students with a first cut of information about a college. Beyond adjusted cutpoints, we strongly feel that input-adjusted *metrics* of any kind are not appropriate to display for consumer purposes.

If the Department, despite these concerns, does choose to make some sort of adjustments, we recommend that it should at least avoid setting different cutpoints for “low,” “medium,” and “high,” and it should maintain a clear comparison to all institutions. Using the current Scorecard design as an example, the comparison point to all institutions should remain – and the cut points should continue to be based on all colleges, but an additional comparison point for similar colleges could be included as a secondary comparison point (see Figure 3).

²⁹ Marks, Jonathan (2013). The Data-Driven Have to Be Strong. Washington, DC: The Chronicle of Higher Education. Retrieved from: http://chronicle.com/blogs/conversation/2013/09/09/the-data-driven-have-to-be-strong/?cid=at&utm_source=at&utm_medium=en.

Figure 3: If an adjusted peer group is used, it should be supplemental to a comparison with all colleges



It is tempting to highly contextualize performance on the metrics through adjusted cutpoints or metrics under the assumption that students will better understand the chances of success for “students like them” at particular institutions (although this may just as likely be interpreted negatively by students as identifying which institutions are “good enough for students like them”). However, the aggregate data available from IPEDS (e.g. SAT/ACT scores, diversity of student body, outcomes by demographics) better describe the profile of institutions, rather than individual students, making it extremely difficult to produce personalized “odds” for students that are accurate enough without a student unit record system. The aggregate information we have on institutions can be utilized, though, to complement the consumer ratings with information on similar but better-performing institutions.

Complement the ratings with information on similar but better-performing institutions

The first tier of information – presented in a way similar to the existing College Scorecard but with easier ability to compare across institutions as described above – provides students with simple, clear, and standardized information that paints an important initial view of institutional performance across colleges of interest to them. However, choosing a college is a complex activity that will require several iterations of research and investigation. A second, dynamic level of the rating system could generate suggested lists of institutions that are similar, but better performing than the college the student first input into the system. These alternate institutions could be output alongside the initial institution of interest on the Scorecard in a table format and/or the “value” grid visual shown in Figure 1.

For this second tier analysis, more refined and manageable peer groupings could be warranted and useful, but must be designed carefully.³⁰ We recommend that the groupings remain fairly broad to avoid limiting students to too few options, and not filter institutions out of the alternate recommendations based on variables like sector that are inconsequential to students. Instead, the groupings could incorporate a broad selectivity measure (e.g. open access vs. non-open access or *Barron's* ratings). Also, this tool will need to consider location when generating a list of alternate options, for three reasons. First, costs at public colleges differ based on the student's state of residence. If alternate school lists are to cross state lines, students will need to input information on their state of residence so the tool can pull in-state cost data for colleges in their state and out-of-state cost information for external institutions.³¹ Second, students often attend college relatively close to home.³² If the tool defines a peer list of institutions too narrowly and without regard for location, all of the schools on the alternate list could be in distant states, providing impractical options for the student. Third, the tool must consider how/when to include online postsecondary options since these cross state lines.

One potential solution to avoid the cost data issues above in the meantime could be to present students with information (again, Scorecard-like data in a table format and/or the "value" graph above) of the 15 top performing institutions within their state, along with a second list of better-performing institutions using a broadly-defined national peer group. The two sets of information would provide students with multiple options, allowing them the opportunity to narrow their choice set further based on their particular needs. This type of dynamic tool could be useful as a second tier ratings analysis, with the potential to broaden students' horizons to include schools they had not previously considered. However, we recommend that the algorithm that generates this list of alternatives be designed carefully and consumer tested.

As noted earlier, the Department could develop an even more dynamic, user-specific system, but only with access to better data. One could imagine a robust online tool that would allow students to enter specific characteristics and interests to generate customized lists of colleges that would serve them well. However, our current data systems are based on simple statistics like means and medians, which do not provide the necessary nuance to generate accurate customized results. For example, IPEDS includes data on the 25th and 75th percentile SAT/ACT scores for many institutions. It could be tempting to compare a student's SAT score to these percentiles and filter institutions for which the student would be

³⁰ The Education Trust's College Results Online (CRO, www.collegeresults.org) includes useful peer groupings of colleges and universities. However, CRO was not originally designed for consumer purposes. Rather, it was developed to help institutions benchmark their performance and set improvement goals and to advocate for institutional improvement by showing the wide ranges of graduation-rate performance among even very similar institutions. The tool's peer grouping methodology is fairly comprehensive, including a whole host of variables – many of which covary. This methodology is appropriate for the tool's original goals, which required fairly narrow peer grouping, but it is not necessarily appropriate for consumer purposes.

³¹ As discussed earlier, net price is a more useful cost measure than tuition and fees or cost of attendance. However, net price is not available for out-of-state students at public institutions. To provide sufficient information to students, the Department should either collect data on the net price for out-of-state students or clearly communicate the difference between cost figures presented for in-state and out-of-state students.

³² Griffith, Amanda and Donna S. Rothstein (2009). *Can't get there from here: The decision to apply to a selective college*, *Economics of Education Review* 28 (2009) 620–628.

in the top or bottom quartile. However, these statistics make clear that a full 50 percent of students at an institution fall outside this interquartile range. Using such a filter could artificially limit choice and hold students back from stretching for reach schools, at which they may be highly successful. A student-unit record system that collected privacy-protected, non-identifiable data could provide, for instance, information on how students in the bottom quartile of SAT/ACT scores perform at different institutions – informing a customized, dynamic college choice tool in a far more precise and effective way.

Invest in counseling: The best consumer information system

The student-oriented aspects of the rating system clearly are well-intentioned. Information – when depicted accurately and simply – can help nudge students towards better choices. However, far too few students, especially first-generation college-goers, have access to the high-touch, data-driven counseling they need to help them interpret college information. In fact, school counselors on average spend only 38 minutes per student per year on college counseling.³³ Even the perfect tool likely will suffer from limited use and effectiveness, unless it is put into the hands of counselors and teachers, who are provided the resources necessary to spend adequate time directly advising students. To take advantage of adults as conduits of college information, we recommend that the Department incorporate these new consumer tools into federal TRIO and GEAR UP programs and train counselors on best practices for use. In the end, data do not counsel people on how to get into college, people do.

Rating System for Accountability

We feel that an accountability rating system must differ notably from a system designed to inform students. While prospective students need simple, unvarnished information, an accountability-based system can allow for a limited amount of additional complexity and nuance. While some guiding principles for an accountability system mirror those of a consumer system, as do the recommended metrics included, we recommend the two be fundamentally different in their design. In addition, while consumer information can be improved now, we suggest that any new system of institutional accountability be phased in over a few years, and institutions have time to adjust to new benchmarks. In designing a rating system intended for the distribution of funds, we recommend that the Department:

- Be clear about purpose and consequences
- Allow for complexity, but strive for simplicity
- Permit only minimal adjustments based on institutional mission
- Set a low, unadjusted performance floor for Title IV eligibility
- Apply graduated penalties using more nuanced mechanisms like a performance matrix, but avoid composite indicators
- Consider a positively-framed incentive-based system

³³ McDonough, Patricia M. (2005). *Counseling and College Counseling in America's High Schools*. Alexandria, VA: National Association for College Admissions Counseling. p. 17. Retrieved from: <http://www.inpathways.net/McDonough%20Report.pdf>.

Be clear about purpose and consequences

While the President’s plan proposes to “tie financial aid to college performance,” it does not specify which federal dollars will be applied or how.³⁴ To make appropriate decisions about the design of a rating system that has real financial consequences, the nature of those consequences must first be made clear. How much money is on the table? What types of financial aid will be tied to performance? If financial-aid dollars are used as the “stick,” how will the accountability system avoid penalizing students? Clear answers to these types of questions will provide a basis on which to make more detailed methodological decisions. We propose a few options for tying dollars to performance in the explanations below. In short, we recommend that the Department set minimum performance floors to determine Title IV eligibility, design a more nuanced mechanism to apply graduated penalties (like a performance matrix), and consider incentive-based structures as well.

Allow for complexity, but strive for simplicity

While a consumer tool must be simple and very easy to understand, an accountability tool has room for more complexity. In fact, making high-stakes funding decisions requires at least some level of complexity to ensure fair, reasoned, and impactful results. However, we recommend that even the accountability-based rating system strive for simplicity to the extent possible. To truly drive improvement, institutions will need to understand how the system works and how they can improve their rating. Certainly well-resourced institutions with robust institutional research (IR) offices will figure out how to respond to, and unfortunately in some instances game,³⁵ even the most complex rating system. However, the less well-resourced institutions with limited IR capacity will be more challenged in trying to understand how and what to improve if the rating system is overly complex. This complexity challenge is particularly problematic because the institutions with limited resources likely are those that the rating system should prod most to improve. For institutions to improve, though, they need a clear understanding of what is expected of them. A “black box” system that includes composites or complex calculations could change the distribution of federal dollars, but will be limited in its ability to change institutional behavior, which is or should be the ultimate aim of the system.³⁶

Permit only minimal adjustments based on institutional mission

With respect to contextualizing institutional performance, there is a lot of debate about whether the Department should only compare institutions to other similar institutions or should adjust for student characteristics. Such comparisons or adjustments may warrant consideration in an accountability

³⁴ The White House (2013). FACT SHEET on the President’s Plan to Make College More Affordable: A Better Bargain for the Middle Class. Retrieved from: <http://www.whitehouse.gov/the-press-office/2013/08/22/fact-sheet-president-s-plan-make-college-more-affordable-better-bargain>

³⁵ Lederman, Doug (2009). ‘Manipulating,’ Er, Influencing ‘U.S. News’. Washington, DC: Inside Higher Ed. Retrieved from: <http://www.insidehighered.com/news/2009/06/03/rankings>.

³⁶ Dougherty, Kevin and Reddy, Vikash (2011). The Impacts of State Performance Funding on Institutions: Research Literature Review and Policy Recommendations. p. 37. New York, NY: Community College Research Center. Retrieved from: <http://ccrc.tc.columbia.edu/media/k2/attachments/impacts-state-funding-higher-education.pdf>

context, although we strongly suggest that the use of “input- or risk-adjusted measures” in particular be weighed very carefully for three important reasons. First, many incorrectly interpret outcomes that have been adjusted for student characteristics as a “ceiling,” thereby lowering expectations for students and institutions. Second, the majority of institutions will mathematically perform “as expected” on adjusted measures, and the more factors included in the statistical model, the fewer institutions that will be identified as over- or under-performing. This would not be a problem, and some would argue that is actually the point, except for the third reason, which reinforces the need for caution. “Expected” or “adjusted” outcomes are calculated based on the current range of institutional performance in the system, meaning that institutions are expected to do as well as the average institution like them. We would argue that, in a system in which fewer than half of students earn degrees on time, there is a lot of room for improvement and that we should not build the status quo into a rating system intended to promote such improvement.

Set a low, unadjusted performance floor for Title IV eligibility

Under current law, institutions are eligible to participate in Title IV financial aid programs if they are accredited by a body recognized by the Department of Education, and they maintain cohort default rates (CDRs) above a bare minimum level.³⁷ In 2011, only eight institutions were subject to sanctions for CDRs exceeding 25 percent in three consecutive years,³⁸ yet in 2009-10, 514 institutions posted default rates higher than their graduation rates.³⁹ Clearly the low bar set by CDRs is insufficient to protect students and taxpayers from poorly performing institutions.

An accountability-focused rating system can serve to strengthen consumer protections and safeguard taxpayer funds by identifying and restricting Title IV eligibility for the truly “worst-of-the-worst” actors. We recommend the ratings use a series of indicators to measure college access, completion, affordability, and post-college outcomes, rather than relying solely on easily-manipulated CDRs (see Section 1 for more details on indicators). Support for the idea of supplementing CDRs with additional measures, or restructuring eligibility for federal dollars is growing. Over the past several years, a number of groups have worked to identify the lowest-performing institutions or recommended strengthening requirements for institutional access to federal financial aid funds, including:

³⁷ Institutions with two-year CDRs greater than 40 percent (starting in 2014, this will change to three-year CDRs greater than 40 percent) in one year lose eligibility to participate in the Direct Loan program. Institutions with two-year CDRs greater than 25 percent in three consecutive years (in 2014, this will change to three-year CDRs greater than 30 percent in three consecutive years) lose eligibility to participate in the Direct Loan and Pell Grant programs. 34 CFR 668.187, 34 CFR 668.206. U.S. Department of Education (2013). Default Rates Continue to Rise for Federal Student Loans [Press Release]. Retrieved from: <http://www.ed.gov/news/press-releases/default-rates-continue-rise-federal-student-loans>

³⁸ Federal Student Aid (2013). Briefing on FT 2011 2-Year National Default Rates. Retrieved from: <http://www.ifap.ed.gov/eannouncements/attachments/2013OfficialFY112YRCDRBriefing.pdf>

³⁹ Gillen, Andrew (2013). In Debt and In the Dark: It’s Time for Better Information on Student Loan Defaults. p. 11. Washington, DC: Education Sector. Retrieved from: <http://www.educationsector.org/publications/debt-and-dark-it%E2%80%99s-time-better-information-student-loan-defaults>

- Education Sector analyzed graduation rates of four-year colleges and universities across the country to identify college “dropout factories,” the approximately 200 institutions nationwide with graduation rates that place them in the bottom 15 percent of all institutions.⁴⁰
- The Education Trust recommended setting minimum performance thresholds on graduation rates, Pell enrollments, and Pell graduation rates to determine eligibility for their proposed federal-state partnership dollars.⁴¹
- HCM Strategists proposed using measures of access, equity, repayment, and risk-adjusted completion rates to strengthen federal aid eligibility requirements.⁴²
- The Department’s various iterations of proposed gainful employment regulations use loan repayment and debt-to-income measures to determine Title IV eligibility for programs intended to prepare students for gainful employment in a recognized occupation.⁴³

We recommend that an accountability rating system— at minimum – identify and eliminate Title IV eligibility for the lowest performing actors, using bright line cutoffs on a core set of measures that are not adjusted for student or institutional characteristics. Because eliminating Title IV eligibility is such a stringent penalty, the bar should be fairly low, targeting only the most harmful institutions. However, all institutions should be held to the same standard, and – regardless of their institutional structure, mission, or student body – be expected to meet a bare-minimum performance floor, such as a 20 percent graduation rate, for example. Any adjustments to this low standard would simply serve to excuse extremely poor performance and downgrade expectations.

Apply graduated penalties using more nuanced mechanisms, but avoid composite indicators

A hard cut for Title IV eligibility based on low thresholds is necessary as a basic consumer protection, but it will not drive improved performance throughout the postsecondary system. While the lowest performing institutions are the most harmful, other institutions certainly could improve on a range of factors – cost, access, success – as well. We feel that a thorough, well-designed rating system should take a multi-faceted approach that incents change among all (or at least most) institutions, while also weeding out the worst actors.

The Department could design this more nuanced rating in a number of ways, each of which carries its own considerations. The technical specifications and design of the system matter immensely and sometimes produce unexpected or unintended results. The Department must test their design and closely evaluate how different institutions perform under various systems to ensure the rating system

⁴⁰ Miller, Ben and Ly, Phuong (2010). College Dropout Factories. Washington, DC: Washington Monthly. Retrieved from: http://www.washingtonmonthly.com/college_guide/feature/college_dropout_factories.php.

⁴¹ Dannenberg, Michael and Voight, Mamie (2013). Doing Away With Debt: Using Existing Resources to Ensure College Affordability for Low and Middle-Income Families. Washington, DC: The Education Trust. p. 12 Retrieved from: <http://www.edtrust.org/dc/publication/doing-away-with-debt-using-existing-resources-to-ensure-college-affordability-for-low>.

⁴² HCM Strategists (2012). Doing Better for More Students. p. 3. Retrieved from: http://hcmstrategists.com/wp-content/themes/hcmstrategists/docs/HCM_Gates_IssueBrief_final.pdf.

⁴³ U.S. Department of Education (2013). Negotiated Rulemaking 2013-14. <http://www2.ed.gov/policy/highered/reg/hearulemaking/2012/gainfulemployment.html>

has the desired impact. Based on our experience and analysis, we offer this review of various ratings options. We strongly recommend against using any composite indicator because such a measure can “wash out” performance on individual measures and create results that are muddled, non-intuitive, difficult to interpret, and generally lack conceptual meaning. After considering the pros and cons of various composite approaches here, we ultimately recommend a matrix-style framework that identifies zones of performance on each metric and considers the interaction between the metrics, which is described in further detail below.

Composite grade system

There are a number of different ways that the Department might construct a composite “grade” for institutions based on their performance on a core set of metrics. For instance, a “star” system – akin to the hotel industry’s 5-star ratings – could operate on a variety of scales (e.g. 3-star, 5-star, 10-star), with a higher star rating representing a “better” school, as measured by performance on a series of measures. Another analog would be an A-F grading system that graded institutions based on combined performance on several measures. While the simplicity and ease of communication of this type of scoring may be desirable, the negative consequences are such that we recommend against using a composite grade system for colleges and universities. Doing so can obscure results as different metrics wash each other out, create indecipherable results among middle performers, and group institutions together within wide ranges of performance.

For illustration purposes only, consider a simple method which applies points to institutions for their performance on graduation rates and net price, in which a combined rating is the sum of each sub-rating. Table 8 shows how three very different institutions can look exactly the same under such a system. In this example, a high graduation rate/high cost institution receives the same rating as a low graduation rate/low cost institution and as an institution that posts mediocre performance on both graduation rates and net price. This is a simple example, and a bigger scale (e.g. 10 points) would allow for more distinction between institutions, but even then, metrics can continue to cancel each other out in ways that may not drive institutional action as intended.

Table 8: Measures can cancel each other out in composite grade system			
Institution	Graduation Rate Score	Net Price Score	Combined Score
A	3	1	4
B	2	2	4
C	1	3	4
3 = Top third, 2 = Middle third, 1 = Bottom third			

This simple example also illustrates how a composite rating can obfuscate results among middle performers, making it difficult to assign consequences to individual institutions. While this type of system is relatively good at identifying top and bottom performers, it creates sometimes indistinguishable results in the middle, as shown by a slightly different example in Table 9.



Table 9: A composite grade system does not distinguish enough between “middle” performers		
Combined Rating	Graduation Rate Rating	Net Price Rating
5 star or ‘A’	High Grad Rate (3)	Low Net Price (3)
4 star or ‘B’	High Grad Rate (3)	Medium Net Price (2)
	Medium Grad Rate (2)	Low Net Price (3)
3 star or ‘C’	High Grad Rate (3)	High Net Price (1)
	Medium Grad Rate (2)	Medium Net Price (2)
	Low Grad Rate (1)	Low Net Price (3)
2 star or ‘D’	Medium Grad Rate (2)	High Net Price (1)
	Low Grad Rate (1)	Medium Net Price (2)
1 star or ‘F’	Low Grad Rate (1)	High Net Price (1)
3 = Top third, 2 = Middle third, 1 = Bottom third		

Ideally, a rating system should reward institutions with high graduation rates and low net prices (5-stars) and penalize those with low graduation rates and high net prices (1-star). But, it does not make sense to treat the three very different, 3-star institutions in the example above the same (see Table 9). In particular, it seems especially problematic to label a low graduation rate institution as a 3-star in this system merely because it is also low cost. This problem can be addressed to an extent by either weighting graduation rates more heavily numerically or trumping cost with graduation rates in this type of scoring system as demonstrated in Table 10.

Table 10. Weighting graduation rates more heavily than net price provides more nuance, but allows for wide range of performance	
10 star	High grad rate & low net price
9 star	High grad rate & medium net price
8 star	High grad rate & high net price
7 star	Medium grad rate & low net price
6 star	Medium grad rate & medium net price
5 star	Medium grad rate & high net price
4 star	Low grad rate & low net price
3 star	Low grad rate & medium net price
2 star	Low grad rate & high net price
1 star	0% grad rate regardless of price

Regardless, a composite grade system based on conceptual cutpoints can create wide ranges of performance within each rating. In our simplified example, the system above would divide graduation rates and net price at four-year institutions into the fairly wide terciles shown in Table 11. Should an institution with a 39 percent graduation rate receive the same boost in the ratings as an institution with

a 55 percent rate? Should a college that has a net price near zero not be rated substantially better than a college that costs \$10,000?

Table 11: Even a 10-star rating includes wide ranges of performance		
Tercile	Graduation Rate	Net Price
Top Tercile	0 – 37%	-\$231 – 10,060
Middle Tercile	38 – 56%	\$10,070 – 13,250
Bottom Tercile	57 – 100%	\$13,260 – 24,674

We feel that composite-based rating systems on their own do not provide enough granularity to adequately distribute federal funds. However, if the Department were to choose this methodology, despite the limitations, we recommend using a bigger scale and using caution about incorporating too many measures. Not only do additional measures make the resulting composite even more opaque, but the inclusion of correlated variables can result in double-counting of outcomes. For example, graduation rates and default rates are negatively correlated, so including both in a composite measure would inadvertently weight completion more heavily than other variables. In short, we recommend that the Department evaluate not just the results of the composite grade in vetting a system of this kind, but also interrogate whether the composite score has sufficient face validity in relation to the sub-scores such that it would actually influence institutional behavior as intended.

Composite Scaled Index

Some of the limitations of a composite grade system could be overcome with a composite scaled index, which might rate each institution from 0-100 using a scaled z-score based on the institution’s z-score on each variable included in the rating system, for example. Under this type of system, different variables could receive different weights (which could also occur under a composite grade system). For example, if the composite were based only on graduation rates and net price, a z-score would be calculated for each institution on each of those variables to convert them to the same scale. These z-scores would then be converted to a 100-point scale and averaged together – either using a simple average or a weighted average, in which one variable receives a higher weight. This approach creates a broader spectrum of performance rankings than the composite grade rating system, but we still recommend against using it for several reasons.

First, the composite score does not have any conceptual meaning besides the fact that a higher score is “better.” In other words, one could not explain what a score of 75 actual *means*. Even the *relative* meaning of how good a score is would change over time based on shifts in the distribution of performance among all institutions. Because the scores essentially measure how far above or below average an institution is, a score of 75 will convey a different level of performance in a year when the mean is 50 than in a year when the mean is 55.

Converting the z-scores into percentile rankings could help with communication and understandability of the metric. However, percentile rankings are just as “relative” as the composite scaled score in so far



as an institution may be performing better than 75 percent of other institutions, but may still be underperforming relative to an absolute benchmark if the range of institutional performance is poor on a given metric (which it unfortunately is on many metrics). Further, whether a z-score or percentile is used, a scaled index has the potential to create “false positives” in so far as the range of performance on the individual metrics can vary considerably at various points along the composite scale since no absolute performance benchmarks are utilized.

Composite Ratio

Instead of rating institutions on measures and adding the resulting scores, or converting performance to z-scores, two measures could be utilized to create a composite using a ratio. For example, a composite measure could be derived by dividing graduation rates by net price.⁴⁴ As with the other composite measures, this type of rating creates a number that does not have any conceptual meaning. If an institution has a 48 percent graduation rate and a \$13,000 net price, the ratio would be 0.37 (using a multiplier of 100), a number that bears little meaning on its own. Like the composite scaled index described above, a ratio also can create “false positives.” However, this methodology also has the added difficulty of creating non-linear output, which makes it difficult to determine at which point an institution moves from “bad” to “good.”

While there are more sophisticated techniques for constructing these types of scales and ratios, our major concern with these approaches is that the gains in precision through the use of more complex methods may come at the expense of institutions being able to clearly understand and act upon the output to improve outcomes for students.

Matrix/Tiered Approach

Instead of using a composite measure to distribute federal funding for higher education, we suggest that the Department consider a mathematically simpler and more transparent matrix approach that incorporates core metrics such as graduation rates, net price, and percent Pell. Each institution would fall in a cell of the matrix based on their performance on those measures and each cell would denote a different level of fiscal penalty or reward as a result. Consider the example matrix in Figure 4.⁴⁵

Utilizing a matrix approach, institutions could, for instance, be required to pay an amount equivalent to a percentage of the Pell dollars they receive into either a risk-sharing fund or direct supplemental grants to Pell Grant recipients.⁴⁶ A similar system might be devised to distribute campus-based aid, set loan

⁴⁴ The Washington Monthly rankings use a similar ratio that divides the difference between an institution’s actual and predicted graduation rate by the institution’s net price. A Note on Methodology: 4-Year Colleges and Universities (2013). Washington Monthly. Retrieved from:

http://www.washingtonmonthly.com/magazine/september_october_2013/features/a_note_on_methodology_4year_co_2046_455.php?page=2

⁴⁵ In this example, the net price cutoffs are set conceptually to represent a net price that a dependent, zero-EFC freshman could finance with (1) subsidized Stafford loans alone and (2) subsidized and unsubsidized Stafford loans. These conceptual cutpoints are logically sound, but in reality, the vast majority of colleges have a net price for low-income students above \$5,500. Different cutpoints could be considered, but it is worth recognizing the level of actual affordability associated with those prices.

⁴⁶ New America Foundation proposed a Pell match for institutions that enroll only small proportions of low-income students, but charge those low-income students high net prices. Burd, Stephen, et.al. (2013). Rebalancing Resources and Incentives in

limits, or scale tax-exempt status, as described further below. The percentage in each cell is determined by an institution's performance on both net price for low-income students and graduation rates as noted in the matrix. However, to "adjust" for performance on an access metric, the percentage could be cut in half if the institution has a higher than average percent Pell compared with other institutions in the same band of graduation rates. If an institution performs in the bottom quartile of graduation rates or has a net price in the highest category, it would not be eligible for the percent Pell "discount," however. Further, to protect against institutions cutting access to improve on the cost and completion metrics, an institution that decreased Pell enrollment in any year after implementation of the rating system could be required to contribute twice the percentage it contributed the previous year.⁴⁷

Figure 4: The Department could use a matrix ratings approach to accountability, in this example, requiring institutions to match the amount of Pell Grant dollars they receive on a sliding scale based on their performance on completion, cost, and access measures.

If % Pell is above average for other institutions in the same column - or in other words, other institutions with similar graduation rates - then the payment shown in the matrix is cut in half.

If an institution decreases its % Pell below what that institution's % Pell (+/- X%) was in the year before implementation of the ratings system, then that institution cannot receive the benefit of moving to a lower-payment box *and* that institution's required payment will double.

		Grad Rate			
		Bottom Quartile (0-31%)	Second Quartile (32-46%)	Third Quartile (47-62%)	Top Quartile (63-100%)
Net Price for Low- Income Students	\$1 - \$3,500 ¹	100%	75%	50%	25%
	\$3,500 - \$5,500 ²	200%	150%	100%	50%
	> \$5,500	300%	225%	150%	75%
		No Title IV			
Average Institutional % Pell ³		61%	52%	43%	30%

¹\$3,500 is the subsidized Stafford borrowing limit for freshmen.

²\$5,500 is the unsubsidized Stafford borrowing limit for freshmen.

³Graduation rate and % Pell numbers are based on Title IV four-year institutions with graduation rate data available in IPEDS.

Hatched cells do not receive a reward for having a high % Pell.

Federal Student Aid. Washington, DC: New America Foundation. Retrieved from:
http://newamerica.net/sites/newamerica.net/files/policydocs/NAF_Rebalancing%20Resources%20FINAL.pdf

⁴⁷ This provision, designed to protect access, could be indexed to the Pell program as a whole and would at least need to include a several percentage point buffer to prevent penalties from impacting institutions with minor swings in percent Pell.

The Department could consider adjustments to this matrix, such as smaller graduation rate or net price bands that offer more distinction between institutions. Predominately bachelor’s institutions could be subject to one set of bands, while predominately associate institutions could be subject to another. However, we recommend that institutions of different control (public, private nonprofit, for-profit) or other distinguishing characteristics (size, Carnegie, selectivity, student demographics) not be subject to different cutpoints for cost and completion on the matrix. The percent Pell measure offers somewhat of an “adjustment” for institutional mission by marking different percent Pell “bonus” standards based on different graduation rate ranges. This approach has the benefit of transparency and clarity around how an institution performs, it attempts to balance accountability for access, success, and affordability, and it avoids the confounding of results that plagues composite indicators.

Options for Leverage: Risk-Sharing, Supplemental Pell, Campus-Based Aid, Loan Limits, Tax-Exempt Status

Different types and levels of consequences could incent different behaviors among institutions. For example, if only small amounts of funding are linked to performance, then institutional behavior may not change. If penalties are dire, then institutions may take drastic steps to ensure compliance. Below, we outline a few leverage point options to consider when designing the accountability rating system. Again, to make reasoned decisions about the rating system design, we recommend that the purpose and consequences first be clearly defined.

Risk-Sharing

The payment determined by the matrix in Figure 4 could be paid into a risk-sharing pot of funds, which could be used to increase Pell Grants. Or, instead of matching the Pell investment in their institution, the matrix could be revised to illustrate how schools could contribute to a risk-sharing fund a portion of all student loan dollars that are in default, not in repayment, or projected to reach forgiveness through income-based repayment. This fund could be used to pay a portion of loan forgiveness or directly decrease students’ outstanding balances.

Supplement Pell Grants

The payments identified in the matrix could also go directly back to needy students in the form of Pell Grant supplements. If an institution has a 100 percent payment requirement, then each student would receive their Pell Grant plus an award from the institution worth the same amount as their Pell Grant. This mechanism, which would function similar to the Pell match proposed by Steve Burd from the New America Foundation,⁴⁸ would have the benefit of directly and immediately lowering costs for low-income students. However, if the funds were to flow directly to students in this manner, each institution’s net price in the following year of data would, by design, decline. Because the result of performance on the matrix directly impacts one of the metrics in the matrix, the Department would need to develop a mechanism for maintaining this system over time.

⁴⁸ Burd, Stephen (2013). Undermining Pell: How Colleges Compete for Wealthy Students and Leave the Low-Income Behind. Washington, DC: New America Foundation. Retrieved from: http://newamerica.net/sites/newamerica.net/files/policydocs/Merit_Aid%20Final.pdf.

Campus-Based Aid

President Obama's FY 2014 budget proposed doling out campus-based aid funds based on institutional success "in enrolling and graduating students from low-income families, setting a responsible tuition policy, and demonstrating good value."⁴⁹ The rating system could tie performance in the matrix framework to these campus-based aid dollars. However, these programs are relatively small in comparison with the larger Pell Grant and student loan programs, so are unlikely to have a dramatic impact on institutional behavior.

Loan Limits

Currently, all institutions are subject to the same loan limits, regardless of how well they prepare students to repay those loans. Using the matrix framework, poorly performing institutions could be subject to lower loan limits, ensuring that students do not accumulate large debts to attend institutions that are unlikely to serve them well. Because loan limits only apply to federal borrowing, the Department should put in place precautions alongside the new federal borrowing limits to avoid a shift towards private loans.

Tax-exempt status

Public and private nonprofit institutions benefit from tax-exempt status to promote their work in the public interest. If they are producing poor student outcomes or failing to serve low-income students, then the level to which they are serving the public good is highly questionable. Instead of being fully tax-exempt, the proportion of the institutional revenues that are tax-exempt could be ratably reduced based on placement in the performance matrix. Eliminating an institution's tax-exempt status would be a severe penalty, akin to losing Title IV eligibility. To avoid such a draconian penalty, institutions could be required to pay some taxes based on a graduated scale. For-profit institutions are not tax-exempt, but to ensure they are captured in an accountability system, the Department could require they pay fines to be calculated in a manner similar to the tax amounts for nonprofit colleges.

Consider a positively-framed incentive-based system

Just as the President has proposed incentive-based plans like the Higher Education Race to the Top and First in the World Fund,⁵⁰ the Department could develop a rating system to distribute incentive funds. The awards would need to be sizeable enough to incent change, and if they are, a competitive effort of this nature could drive a variety of improvements. Again, the consequences associated with the rating system (in this case, positive consequences) matter deeply when designing it. For an incentives-based system, the Department could apply a "points system" to divide up a set amount of funds among institutions similar to performance-based systems currently being utilized to distribute appropriations in a growing number of states. Colleges could accumulate points on a variety of student-based measures, such as the number of students who persist to year two, earn 24 (or 30) credits in the first two years, or

⁴⁹ Department of Education (2013). The Budget for Fiscal Year 2014. p. 83. Retrieved from <http://www.whitehouse.gov/sites/default/files/omb/budget/fy2014/assets/education.pdf>.

⁵⁰ Department of Education (2013). The Budget for Fiscal Year 2014. p. 82-83. Retrieved from <http://www.whitehouse.gov/sites/default/files/omb/budget/fy2014/assets/education.pdf>.

earn a credential. Institutions could receive extra points for ensuring students of particular national interest, such as Pell Grant recipients, meet these momentum points.⁵¹

The matrix approach could be adapted to include incentive funds in some of the performance cells and/or to reward institutions that move between cells within the matrix, which we encourage the Department to consider if the goal of the accountability rating system is broad enough to encompass both rewarding and sanctioning institutions. However, if the Department is more narrowly aiming to devise a method for distributing competitive funds (or campus-based aid funds as some have suggested), a points-based system may be more straightforward. A points system is less well-suited to broad-based accountability for Title IV funds, however, since those funds are awarded to students not institutions like state appropriations are in the PBF context from which this system is drawn. If a points system were to be adapted to determine Title IV related sanctions and rewards across the board, the Department would experience some of the same issues affecting the composite grade or scale systems in so far as determining a meaningful level of performance (e.g. number of points) either in a given year or over time. This raises the need to address the important issues raised below with respect to clearly defining what the President and the Department *means* by tying financial aid to college performance.

4.0 Presentation of Ratings Information

If the purpose of the rating system for students is to shape student behavior in ways that make them more likely to enroll in college, and enroll in colleges that will serve them well, the construction and presentation of the ratings must be carefully considered, as was addressed in previous sections. Despite the inherent challenges in the task, it is an important one that research indicates could yield real results for students. Students – particularly those from traditionally underserved populations – who receive clear information on college costs and financial aid options, are more likely to take steps toward attending college, particularly four-year institutions.⁵² And, providing high-achieving, low-income students, in particular, with timely and semi-customized information on the college application process, as well as net costs, increases their likelihood of applying to and attending selective colleges, which tend to have higher graduation rates.⁵³

Policymakers have not been blind to the need for better information in the hands of students. In the 2008 reauthorization of the Higher Education Act, Congress mandated the eventual use of net price calculators by schools, which are now in place. Since then, several federal initiatives, including the College Scorecard and the Financial Aid Shopping Sheet, have also been developed to provide standardized, clear information to students to help compare institutions. In the case of the College

⁵¹ Value-Added Funding. Complete College America. Retrieved from: <http://dl.dropboxusercontent.com/u/28697036/Value-Added%20Funding%20Do%20This.pdf>

⁵² Berkner, L., & Chavez, L. (1997). *Access to Postsecondary Education for 1992 High School Graduates*. Washington DC: National Center for Education Statistics. p. 59-61. Retrieved from: <http://nces.ed.gov/pubs98/98105.pdf>

⁵³ Turner, S. and Hoxby, C. (2013). Expanding College Opportunities for High-Achieving, Low Income Students. *Stanford Institute for Economic Policy Research*. P. 23-26. Retrieved from: <http://siepr.stanford.edu/?q=/system/files/shared/pubs/papers/12-014paper.pdf>

Scorecard and Shopping Sheet, rigorous evidence does not yet exist on the overall impact of these tools on college awareness, attendance, and outcomes. However, students, higher education experts, and policymakers have expressed concern that some information on these tools (such as net price, borrowing and repayment options) requires a level of financial literacy that does not exist among many families, particularly low-income families.⁵⁴ The U.S. Department of Education has acknowledged these potential hurdles – and taken action to address them – by recently announcing updates to the voluntary Shopping Sheet, including a glossary to help explain financial aid terms to students.⁵⁵ In developing the consumer rating system, we recommend that the Department carefully consider which metrics require more explanation and contextualization for students, and be transparent about why a metric is used at all to rate colleges. Further, there should be extensive focus group testing of the system, particularly with nontraditional student populations.

Likewise, as previously discussed, we recommend that the accountability rating system include metrics that are easily understood, and that are presented in ways that can easily shift institutional behavior in beneficial directions. A “black box” approach makes it more difficult for institutions to determine how they could improve, which is the goal of the system.

5.0 Existing Rating Systems

Consumer Information Ratings and Metrics

In modeling rating systems, the Department could look to several consumer-focused initiatives outside of higher education that have proven successful at influencing both consumer and industry behavior. Ratings are used widely by consumers, businesses, and the public sector to differentiate and gain more information for products and/or services. The level of complexity ranges from scale ratings (one to five stars), binary ratings (such as a thumbs up or a thumbs down), to more complex algorithms you may find in specific industries such as rating systems used in banking. Some ratings, such as the Consumer Reports Automotive Ratings,⁵⁶ conduct extensive research surveying hundreds of thousands of car owners and conducting their own tests to provide consumers with a tiered rating system. Over the past few decades, social rating systems⁵⁷ have grown in popularity, allowing users to collectively evaluate the quality of an item, and often assign a binary or star rating, and increasingly include reviews in the assessments.

⁵⁴ NASFAA and JBL Associates (2012). *No Clear Winner: Consumer Testing of Financial Aid Award Letters*. Washington, DC: National Association of Student Financial Aid Administrators. P. 19. Retrieved from: http://www.nasfaa.org/advocacy/award-letter/No_Clear_Winner_Consumer_Testing_of_Financial_Aid_Award_Letters_-_Summary_Report.aspx. Additionally, in preparation for various forthcoming reports, IHEP held approximately a dozen meetings with higher education experts, former and current staff from the White House, Department of Education, the House and Senate. The concern about adequate financial literacy was raised during many of these meetings.

⁵⁵ U.S. Department of Education (2013). Financial Aid Shopping Sheet Updated to Provide Students with More Transparency. Retrieved from: <http://www.ed.gov/blog/2013/12/financial-aid-shopping-sheet-updated-to-provide-students-with-more-transparency/>

⁵⁶ How Consumer Reports Automotive Ratings Work, Sarah Winkler, How Stuff Works <http://auto.howstuffworks.com/buying-selling/consumer-reports-automotive-ratings4.htm>

⁵⁷ 7 Things You Need to Know about Social Rating Systems: <http://ets.tlt.psu.edu/wp-content/uploads/socialratings1.pdf>

In the case of *Consumer Reports*, consumers can compare a product by type (for example, a front-loading washing machine vs. a top-loading washing machine), as well as view products ranked on a series of metrics (for example, energy efficiency) or on a composite score. As we noted previously, we recommend that the Department consider disaggregating the ratings on each metric rather than simply displaying a composite rating score. Doing so will allow consumers the opportunity to take away the information that is most important to them, as well as allow institutions to see where best they can improve.

Other rating systems have been created or backed by the government to persuade consumer behavior. One example is Energy Star,⁵⁸ a voluntary program from the Environmental Protection Agency (EPA) in which products (or homes) are given a certification that they meet basic levels of energy efficiency. Energy Star could inform rating systems that identify institutions that are producing a minimum level of quality for students, as well as identify the bottom performers among institutions on a series of metrics.

A final example is automobile fuel economy.⁵⁹ In this case, car companies test the fuel efficiency of vehicles and report results to the EPA (who may audit 10 to 15 percent of the results to ensure accurate reporting). The EPA, in turn, provides information on Miles per Gallon (MPG) for every make and model of car distributed in the U.S., and requires this information to be affixed to every new passenger car and truck. Recent research indicates that consumers broadly support higher mileage standards, and consumer behavior has shifted in favor of more fuel-efficient cars.⁶⁰ Again, for the rating systems, the embrace of MPG by the car-buying public should give the Department confidence that with improved data, consumers may understand and respond to certain higher education metrics – say, graduation rates – if they are easily understood. Further, like MPG, students may need time and assistance putting higher education metrics in context. We recommend that the Department pair any metrics with clear and digestible information on what exactly the metric means, and how they can be interpreted.

⁵⁸ <http://www.energystar.gov/>

⁵⁹ <http://www.fueleconomy.gov/>

⁶⁰ McGillis, J. and Cooper, M. (2013). On the Road to 54 MPG: A Progress Report on Achievability. Consumer Federation of America, <http://www.consumerfed.org/pdfs/ON-THE-ROAD-TO-54-MPG.pdf>