December 2, 2013

Jamie Studley
Deputy Under Secretary
Office of the Under Secretary, Department of Education
400 Maryland Avenue, SW, Room 7E300
Washington, DC 20202

Dear Deputy Under Secretary Studley:

The Association of American Universities, a nonprofit association of 60 U.S. and two Canadian preeminent public and private research universities, would like to respond to the President’s “Plan to Make College More Affordable: A Better Bargain for the Middle Class,” released on August 22, 2013. AAU applauds the President’s focus on college affordability at a time when Americans are concerned about the cost of a college education, the impact of student loan debt, and future employment prospects. AAU member institutions strive to provide the highest quality education possible and to make it affordable through a range of cost-savings measures, substantial institutional financial aid programs for low- and middle-income students, and creative use of technology. AAU is encouraged by many of the goals outlined in the three sections of the President’s plan: paying for performance, promoting innovation and competition, and ensuring that student debt remains affordable. This response is focused on the “Paying for Performance” section of the plan, which calls for a new college rating system by 2015, with a link to the awarding of federal student aid by 2018.

AAU does not endorse a new ratings system, but we would like to share our views. As articulated in prior statements about the Administration’s consumer disclosure proposals, we believe that any tools designed to be useful to students and parents in their college search should be grounded in data that are reliable and valid, and that are presented with the appropriate context to accurately reflect institutional performance. In preparing this response, we asked AAU institutional research officers and other campus experts to provide us with feedback on the reliability and validity, as well as methodological limitations, of possible metrics related to access, affordability, and outcomes. The sheer volume of responses from these experts speaks to the inherent complications and controversies for this—or any—rating system. We have summarized the responses below and divided the responses into (1) general feedback on rating
systems, (2) data sources and definitional issues, and (3) peer groups, and we have also provided specific feedback on metrics linked to access, affordability, and outcomes.

**General Feedback**

*Limitations of Rating Systems*

Despite the proliferation of college rating and ranking systems over the last several years, developing a valid and useful rating system is a daunting endeavor. Given the diversity of higher educational institutions (including their varying sizes, locations, and missions), the shortcomings of available data, and the breadth of the Administration’s stated policy objectives, the exercise becomes even more controversial and complex. While the Department of Education website already contains a great deal of information about institutions drawn from a variety of sources, combining a particular set of measures into a single rating raises important questions, including: What methodology will be used to generate the ratings? What measures will be used? How will different measures be combined and weighted? What are the appropriate peer groups for comparison? Do these variables (methodology, measures, weights, peer groups) vary for different objectives? Depending on the answers to these questions, a system may be biased against some group(s) of institutions and toward others, leading to unintended consequences for students and institutions.

One central dilemma, especially if such a rating system is intended for parents and prospective students, is that many of the available measures and metrics require additional context to be useful. Yet providing more context for individual measures seems in opposition to the approach of combining measures into a single rating. Different types of institutions have different missions and may require different metrics to capture their strengths and weaknesses. A one-size-fits-all approach fails to capture the diversity of the U.S. higher education system.

A further difficulty is that a rating system that looks at access, affordability, and outcomes will of necessity consider both the inputs and outputs of higher education. Yet these are not independent of one another: designing a rating system that accounts for these relationships in a way that does not provide incentives to institutions to engage in undesirable behaviors will be a major challenge.

*Data Sources and Definitional Issues*

Even if a perfect methodology for developing a rating system existed, it would be meaningless without good data. As described, the President’s rating system — like other resources for students on the Department of Education website, such as College Scorecard and College Navigator — would likely draw from extant data sources, primarily the Integrated Postsecondary Education Data System (IPEDS). While IPEDS was initially developed to monitor trends in higher education, over the years it has become a tool for enforcing institutional compliance and assessing institutional effectiveness, with mixed results. Fundamental limitations are built into the information in IPEDS, which is self-reported by institutions. Below we provide one example
where schools under-report information, and other examples where the data are available only for a subset (and perhaps not always a representative subset) of students. A rating system must be equipped to deal with missing or limited data, and the use of inappropriate denominators to calculate ratios and percentages is a concern in the construction of many metrics based on student data. Overall, while IPEDS data can be informative for parents and students, they may not provide the whole picture, and require contextualization to be most useful. IPEDS data are inappropriate as the basis for a system that will ultimately make determinations about institutions’ eligibility for federal financial aid.

Vague definitions of data specifications or policy terms can also be a problem. Often, it is the availability of extant data that determines the definition of policy terms such as access, rather than stakeholders first agreeing on definitions, with the search for the right data following suit. Such an approach may lead to measuring what’s available, rather than what stakeholders truly care about.

Regardless of the final data sources for a rating system, the department should ensure that all data are publicly available and made available to colleges and universities for review prior to publication of such a rating system.

Peer Groups
There is much concern among AAU institutional researchers about how peer groups would be defined in a rating system. Higher education institutions are nearly as diverse as they are numerous, and systems to categorize institutions inevitably have shortcomings. While groupings like the Carnegie Classification capture many of an institution’s core characteristics, no classification system captures features such as whether the institution is committed to need-blind admissions or to providing institutional grant aid to meet full demonstrated need of eligible students.

A related issue involves how performance within a rating system will be measured within peer groups: will institutions be compared to one another, to themselves over time, or to an external benchmark of some kind? We caution that comparisons within any peer group necessarily lead to “winners” and “losers.” Such an approach seems inconsistent with the Administration’s goals of encouraging all institutions to innovate and improve their performance.

Specific Feedback
Beyond the more general issues outlined above, we wish to provide comments on specific types of information that may have a role in a rating system, some of which are currently part of the College Scorecard.
Cost and Access: Currently, the College Scorecard reports average net price and two-year increases in average net price. We believe that access and affordability would be better represented by utilizing average net price by income, as currently reported through IPEDS. We also note that cost at public universities is tied closely to state appropriations and other policies; a rating system should take account of these differences so as not to penalize institutions from particular states.

Numbers and percentages of Pell grant recipients are a simple and direct way to look at access, yet these figures should be compared to the overall population of Pell-eligible students, not necessarily all undergraduates. Looking at Pell recipients also provides a narrow view of access, in that it potentially ignores a large population of low-income, middle-income, and first-generation students who should be part of any access equation. Many students who do not qualify for Pell grants still face considerable financial need. Information on institutional aid, collected through IPEDS, should also be included, as it demonstrates institutional commitment to access and affordability. Finally, some contextual information (e.g., comparisons to state or regional demographics) would give Pell grant figures more meaning.

Graduation rates: The graduation rate reported on the College Scorecard is the six-year graduation rate for first-time, full-time students, which makes sense for some institutions but may not make sense for others. We are also concerned about any potential use of transfer-out rates: in 2011, only 27% of four-year institutions reported transfer-out students to IPEDS. Additionally, the six-year graduation rate as currently reported does not account for students who transfer from one institution and complete their degree at another; other efforts, such as the Student Achievement Measure initiative, contain more comprehensive information about student progress and graduation. Some thought also must be given to institutions that are part of multiple-institution systems.

Regardless of the rating system, we continue to encourage the Department of Education to make progress in improving existing data sources such as the graduation rate data in IPEDS. The need for better data arguably trumps the need for another college rating system.

Loan default rates: Loan default rates are important, but we believe that further honing the information used would make it more useful for students. Loan default rates matched with a particular institution should apply only to debt accrued at that institution. Further, a separate loan default rate should be calculated for graduates only, as this would provide a better point-of-comparison for prospective students. Finally, to properly contextualize this figure, the number of students in default and the number of graduates should be reported, in addition to the default rate.

Median borrowing: Like loan default rates, borrowing and debt are important data, but they require honing and contextualization to be most useful. Borrowing linked with a particular

1 http://www.studentachievementmeasure.org/
institution should apply only to borrowing that students accrued at that institution. Currently, median borrowing for an institution includes both graduates and non-graduates who entered repayment; these groups should be separated to provide better points-of-comparison for prospective students. We also believe that, in addition to providing a median figure, the number and percentage of graduates who have no debt should be reported. The figures as currently reported on the College Scorecard imply that they apply to all students, not just those who borrowed.

**Employment:** We believe that this section has great potential to do more harm than good. In particular, it is hard to understand how reporting one earnings figure for each institution could be of any use to parents and prospective students without somehow properly accounting for: 1) the mix of majors at the institution, 2) the location of most graduates (e.g., cost-of-living differences), and 3) the timeframe after graduation at which earnings are considered. Reliable data on college earnings demonstrate that the greatest variation in graduates’ earnings is within institutions, rather than across institutions, with these differences being driven by choice of major and occupation, prior employment experience before or during college, and other issues.

Beyond these three factors, there are additional complications. Certain occupations have high societal value, yet pay relatively low salaries (e.g., teachers, social workers); this must be taken into account so as not to provide a disincentive for institutions to provide degrees in these fields, and for students to enter them. Graduates who go on to attend graduate or professional school may earn very low wages during their course of study, and sometimes beyond (e.g., in the case of postdoctoral researchers in some STEM fields). The methodology used by Payscale.com bypasses this issue by excluding students who go on to earn additional degrees, but using such a methodology (i.e. restricting earnings to only those who earn terminal bachelor’s degrees) would present a biased, and inaccurate, picture of outcomes for many of our institutions. Even in higher-paying fields, rather than comparing salaries across fields or institutions, a more useful exercise would be to look at graduates’ salaries relative to others in the same field, location, and career stage.

The earnings data available for this type of accountability measure also have severe limitations. Even if the Department of Education is able to link Social Security or Unemployment Insurance (UI) wage data of individuals back to the colleges they attended, there would still be other issues that must be addressed. Many students today attend more than one college, so it can be difficult to apportion the value of all of their earnings to any single institution. The wages of individuals who are working less than full time, either by choice (due to parental and other caretaking responsibilities, or for other reasons) or due to labor market dynamics, can skew the results for any single institution. In addition, UI wage data do not reflect self-employment income. Using IRS data can help address this weakness, but it has its own limitations because of the difficulty of disaggregating income data for married joint filers if the data used are restricted to Forms 1040/1040A/1040EZ. Finally, linking earnings data to the CDR cohort limits the reporting of that information only to those who have taken out federal student loans. This potentially
excludes a significant number of graduates and may possibly bias the reported figures, particularly for institutions that graduate a majority of their students without debt. It also may include individuals who did not graduate, which is not useful for prospective students.

Finally, it goes without saying that the value of a college degree goes beyond maximizing earnings. Yet a ratings approach risks quantifying a complex variable like earnings in a way that fails to provide valuable information to prospective students, while simultaneously ignoring and excluding the non-quantifiable ways a college degree contributes toward pursuit of a rewarding and satisfying career. A primary focus on earnings as a measure of institutional performance also exacerbates the recent trend of focusing almost exclusively on the private returns to higher education. It unwisely diminishes the social value that a college graduate provides to his or her community and our nation.

Thank you in advance for your consideration of AAU’s views. We look forward to a conversation in the coming months about how we can enhance college access and promote college completion while protecting the integrity of federal student aid funding and encouraging states to reinvest in public higher education.

Sincerely,

[Signature]

Hunter R. Rawlings III
President
Association of American Universities