A New System of College Ratings– Invitation to Comment

I. Introduction

A. Background

Since taking office in 2009, President Obama has set a course for the nation to once again lead the world in college attainment as it did a generation ago. To help us meet this goal means ensuring all Americans have the opportunity to enroll and succeed in college. Our nation's success in achieving this goal leads us to greater economic prosperity with a thriving middle class and a stronger, more secure democracy.

To date, the Administration has taken historic steps towards achieving this goal. We have worked with Congress where possible to make college more affordable for working and middle class families, including increasing the maximum Pell grant award by nearly $1000, creating the American Opportunity Tax Credit worth up to $10,000 over four years of college, enacting student loan reforms that eliminated bank subsidies and reinvested dollars back into student aid for American families, and keeping student loan interest rates low. The Administration has taken steps to ensure that no student has to spend more than 10 percent of his or her monthly income to repay their federal student loans and to protect students from unaffordable debts resulting from poor-quality career college programs through tougher Gainful Employment regulations.

Most recently, the President, Vice President, and First Lady also issued a call to action on college opportunity in partnership with college and university presidents, superintendents, high school and college access counselors, business, and philanthropy to announce over 600 new actions to help more students prepare for and graduate from college by building networks of colleges focused on promoting more degree and credential completion, creating K-16 partnerships around college readiness, investing in high school counselors as part of the First Lady's Reach Higher initiative, and increasing the number of college graduates in STEM fields (science, technology, engineering, and mathematics).

Today, almost six years later, there are many promising signs of progress towards the President's goal, including record high-school graduation rates and increased college enrollment, notably among Hispanic and African Americans. And while much progress has been made, there is more work to be done. Too many American families still feel that a college education is out of reach. Too many students find that they need to take remedial coursework before they can even begin courses that lead to a degree. Too many students do not complete their education or find that their education does not prepare them well for work in their chosen field or to repay their educational debt. Too many states have reduced their investments in higher education, leaving students and families shouldering more of the burden to pay for college.

Last year the President sought to tackle these challenges by announcing an ambitious agenda to combat rising college costs, make college affordable for American families, and strengthen the value of a college education. A key part of that effort has been the Department of Education's work to strengthen the performance of colleges and universities in promoting access, ensuring affordability, and improving student outcomes through the design of the college ratings system. The system aims to: (1) help colleges and universities measure, benchmark, and continue to improve across the shared principles of access, affordability, and outcomes; (2) help students and families make informed choices about searching for and selecting a college; and (3) enable the incentives and accountability structure in the federal student aid program to be properly aligned to these key principles.

As the Department progresses in its efforts to design a ratings system that is clear, fair, and focused on a few key critical measures of institutional performance, while accounting for the diversity and complexity of the nation's rich system of higher education, it will continue to refine the ratings system over time based on user and institutional experience, input from the field, and the availability of additional data.

Additional information and questions relating to the specific metrics under consideration and how they might fit together are included in more detail below. It is important to clarify that this document is not version 1.0 of the ratings. As mentioned previously, version 1.0 will be released prior to the 2015/2016 school year. This document provides a bigger picture framework that will further guide the development of version 1.0.

B. Characteristics and Purposes of the Ratings System

The college ratings system has multiple related purposes. A critical purpose of the ratings system is to recognize institutions that are succeeding at expanding access, maintaining affordability, and ensuring strong student outcomes and setting them apart from institutions that need to improve. By shedding light on key measures, the ratings system will support greater accountability and incentivize schools to make greater progress in these areas of shared priorities, especially at serving and graduating low income and first generation students and holding down the cost of college. In and of itself, expanded information around a small number of critical measures, with appropriate attention to institutional variation, can be useful to the general public, researchers, philanthropists, policymakers, and other decision-makers.

Researchers and policymakers will have new tools to examine colleges' performance, identifying underperforming institutions, and recognizing excellence, which can in turn help the academic enterprise explore the policies and practices that can strengthen efforts at colleges and universities to improve. Some existing college "rankings" have rewarded exclusivity and rising college costs, moving our nation away from tackling some of the toughest challenges facing our American higher education system. Using better performance information (and continuing to improve upon the data we have) can help realign the incentives towards our goal of helping more students from all backgrounds succeed.

In the words of President Joseph M. Sopcich of Johnson County Community College (Overland Park, KS), "We use data to make ourselves better." Colleges will have the ability to view their own performance and compare themselves to their peers on measures of access, affordability, and outcomes—enabling them to better identify their own areas of strength and weakness—and identify other institutions to learn from and partner with. This type of benchmarking will facilitate continuous improvement and help colleges to better target their efforts and resources as they seek to improve and become more competitive.

A second critical purpose is to help students and families make more informed choices during the college search and selection process. By offering clear, accessible and reliable information about critical measures of access, affordability, and outcomes the Department believes that the ratings system will begin to inform consumer behavior—over time leading to better educational decisions and, ultimately, improved student outcomes. While the Administration has taken steps to improve consumer information through efforts like the College Scorecard, we need better information, tools, and resources to help students and families select schools that provide good outcomes at an affordable price. Without this, there is insufficient incentive within the higher education marketplace to offer a quality education at a lower price, or for low-performing institutions to improve.

Finally, the Federal Government has a unique responsibility to both students and taxpayers to ensure wise and effective use of the $150 billion invested in the federal student aid program annually. For that reason, as the ratings system is refined, the Administration will consider how to use these newly available data to invest federal student aid where it will do the most good.

C. Overview of Outreach Activities

From the outset, President Obama and Secretary Duncan committed to consult widely to develop the college ratings system with broad public input from key stakeholders and diverse perspectives through a combination of open meetings, an invitation for public comment, and consultation. On September 19, 2013, the Department launched the listening process with a meeting with national student leaders. During Fall 2013 the Department held open forums in Los Angeles, California (California State University-Dominguez Hills), Cedar Falls, Iowa (University of Northern Iowa), Baton Rouge, Louisiana (Louisiana State University), and metropolitan Washington, DC (George Mason University). Throughout the year, department officials met with institutions and organizations across the country, including in Boston, Richmond, Annapolis, Chicago, Kansas City, Denver, Boise, Las Vegas, San Diego, and Davis (CA), to discuss the ratings in small and large sessions.

To ensure that a broad range of stakeholders could give input, in September 2013 the Department also established an email address (collegefeedback@ed.gov) for any interested person or organization to submit ideas. To date we have received more than 175 comments through that email address from a wide range of students, parents, college presidents and other leaders, faculty, counselors, and others.

In December 2013, the Department published a Request for Information (RFI) in the Federal Register to gather technical expertise pertaining to data elements, metrics, data collection, weighting, scoring, and the presentation of the ratings system. We received more than 140 comments from a range of individuals, organizations, researchers, and institutions.

The National Center for Education Statistics convened a technical symposium in February 2014 to discuss ratings measures, systems, and challenges. The day-long event featured some of the most prominent experts in these areas in higher education as well as some experienced with similar issues in the health care field.

During June andJuly, the Department conducted telephone consultations with more than three dozen academic and policy experts familiar with postsecondary data sources; institutional, state, voluntary, and other ratings and performance methodologies; and specific categories of institutions.

Overall, the outreach process has generated more than 160 sessions, meetings and calls across the country with more than 9,000 students, parents, and counselors; college trustees, presidents, faculty, and administrators; members of Congress and state and local leaders; education researchers, ratings experts, and policy analysts; and student advocates. The outreach process has included presentations and discussions by senior leaders with a wide range of national and state education and policy organizations and diverse institutional types. Special sessions and outreach were arranged with community colleges; minority-serving institutions; rural and private colleges; and families with little college-going experience and the people who advise them. Comments have been thoughtful, constructive, extensive, and, as the discussion proceeded, increasingly specific about particular metrics, incentives, risks, alternative methodologies, and data resources. Many of the comments are reflected in the directions and questions that follow.

II. Overview of Ratings System

As the President laid out in his August 2013 remarks, the ratings system will be based on a set of measures that address the main domains critical to consumers, institutional leaders, and policymakers:

- *Access,* such as percentage of students receiving Pell grants;
- *Affordability*, such as average tuition, scholarships, and loan debt; and
- *Outcomes*, such as graduation and transfer rates, graduate earnings, and advanced degrees of college graduates.

The Department has been working to determine metrics and an overall framework that will effectively advance that vision. Additional information and questions relating to the specific metrics under consideration and how they might fit together are included in more detail below.

A. How the System Will Rate Institutions

Rating Categories: From the outset, the plan has been to create a college *ratings* system, using broad categories to highlight significant success and serious weakness, and not a numerical *ranking of institutions.* The Department has received many comments that the ratings should be focused, clear, and simple, and should avoid "rankings" and false precision. The Department plans to limit the ratings to a small number of performance categories. The Department is considering only three levels for each category: high-performing, low-performing, and those falling in the middle. This will allow the ratings system to identify institutions that are performing very well, and those that are weak or struggling, without suggesting more nuanced distinctions than the data can meaningfully support. For consumers, policymakers, and the public, the Department feels that identifying extremely high- and low-performing institutions using a simple set of key measures is the most valuable action the Department can take. The Department is currently conducting data analyses to determine the best way to set the thresholds for what constitutes high or low performance.

Improvement over time: Another important feature of the ratings system is whether and how to recognize performance improvement over time. The Department has received very positive feedback on the concept of recognizing improvement. The Department appreciates that an important purpose of the ratings system can be to encourage and support institutional improvement. The Department is interested in exploring how the system should give recognition or credit to institutions that are demonstrating meaningful improvement. We believe presenting this information in addition to an institution's rating(s) will provide the public with useful information not only about the institution's current performance, but also about how its performance is trending over time. We are exploring ways to incorporate improvement into the system. Specifically, the Department is interested in comments on the following questions:

- *Which metrics should the Department look at to identify institutional trends over time?*
- *Over how many years should the Department look to identify institutional trends over time?*

Institutions that will be included in the ratings system: The Department has received thoughtful feedback on the implications of determining which institutions are included in the ratings system, and which students within those institutions should contribute to an institution's rating. In the first version of the ratings, the Department is planning to include:

- Predominantly four-year institutions, which are colleges that primarily award baccalaureate degrees and above; and
- Predominantly two-year institutions, which are degree-granting colleges that primarily award associate degrees and/or certificates.

The predominant award level of the institution will be based on the number and level of awards that the institution reports on the completions component in IPEDS. For instance, if the majority of degrees/certificates awarded are baccalaureate degrees, then the institution will be classified as a predominantly four-year institution. Likewise, if an institution is comprised of 30 percent certificate programs, 30 percent associate degree programs, and 40 percent baccalaureate degree programs, the institution would be classified as a predominantly four-year institution.

At this time, both graduate-degree only and non-degree granting institutions will not be included in the first version of the ratings. For institutions that serve graduate students in addition to undergraduates, we will only take the undergraduate student population into account.

The Department has concerns on the data limitations for non-degree granting institutions. For example, a completion rate may reflect the length of a program since certificate programs are as short as a semester, while other certificate programs are closer to two years in length. This will make it difficult to make fair comparisons between non-degree granting institutions since completion will be easier to attain at a short-term certificate program, but more difficult to attain at a lengthier program. Another limitation is that virtually all non-degree-granting institutions are "program reporters," meaning they only report net price to IPEDS for their largest program rather than the entire institution. In the future, the Department will have complete data on program length for institutions that operate within this sector, which will allow for a more rigorous evaluation under any future ratings framework. In the meantime, the accountability framework promulgated through the Gainful Employment regulations will help students attending these schools avoid low-performing programs and require disclosures that inform students of a program's performance before they enroll. Even with non-degree granting institutions excluded, the first version of ratings will cover degree-granting institutions that represent about 95% of all undergraduate enrollment.

Institutional Groupings: The Department has also received thoughtful feedback about differences in institutional characteristics and mission, and how those differences ought to be taken into account when rating institutions. Consistent with this feedback and the President's goal to compare schools with similar missions, the Department is exploring several strategies for grouping institutions together for comparison purposes.

The Department is planning on grouping predominantly two-year institutions separately from predominantly four-year institutions, due to differences in missions. For instance, a community college will typically be rated against other predominantly two-year institutions. If the Department later chooses to also include non-degree-granting institutions, they would be grouped separately. This treatment would allow the use of differing metrics if appropriate for different types of institutions.

In addition to presenting four-year institutions separately from two-year institutions, the Department is considering accounting for differences in institutional characteristics such as degree and program mix and selectivity. The Department is currently conducting data analyses to determine what, if any, more granular groupings might be used, and welcomes comment on this topic.

B. Data

This first version of college ratings will rely primarily on data from federal administrative data systems or existing federal data collections. The most important sources include the Integrated Postsecondary Education Data System (IPEDS) and the National Student Loan Data System (NSLDS).

Through IPEDS, all institutions that participate in federal student financial aid programs are required to report data on enrollments, program completions, graduation rates, faculty and staff, finances, institutional prices, and student financial aid to the National Center for Education Statistics (NCES) on an annual basis. Institutions report data to IPEDS on an aggregate level, with limited data available for some subpopulations of interest (e.g., net price disaggregated by income quintiles).

In contrast to the institution-level data available in IPEDS, NSLDS is an administrative data system that contains individual-level data on students who receive federal student aid. NSLDS includes student financial aid application data including demographic and socioeconomic status information reported on the Free Application for Federal Student Aid (FAFSA), loan repayment data, and institution-specific enrollment data such as transfers and completion. Unlike IPEDS data, NSLDS data cover only students who receive federal student aid.

As explained, the Department will be using existing sources of data for the first version of the ratings. These are data that have already been reported by institutions or are data reported by an individual, such as through the FAFSA. Going forward, the Department would like to support institutions to ensure that data are accurately reported and that the data underlying the calculated metrics reflect the best available information from the institution. The Department is exploring the feasibility of a simple process for institutions to review their reported data. In any event, the Department strongly encourages institutions to ensure that the data they have reported are accurate.

The Department is also considering allowing institutions to provide alternative sources of data to supplement their information in the ratings system. In addition to the improved completion measure in IPEDS and the alternative measure of completion that we are considering developing using data from NSLDS, the Department is exploring external measures validated by use in the field to better capture completion outcomes for a broader range of students on a campus. For example, in a future iteration of college ratings, the Department may allow institutions to voluntarily submit their Student Achievement Measure (SAM) completion/transfer rate or possibly to consent to the use of institutional data held by other federal or state agencies for calculation of specific outcome metrics for a broader cohort of students.

In addition to releasing the ratings, the Department will release the data used to construct them to facilitate transparency and provide policymakers and researchers with data to inform their efforts to identify which policies best increase access and improve quality.

All data used for the ratings will be used and maintained according to the highest federal privacy standards. No student-level data will ever be released to the public; all student data will be presented only in aggregate.

C. Metrics

The Department has been investigating and securing input on possible metrics to illuminate the key areas of access, affordability, and outcomes. Rating colleges based partly on the access that they provide to disadvantaged students gives institutions an incentive to enroll such students in greater proportions. Additionally, our goal is to rate colleges on affordability based on the costs of attendance that students incur after subtracting out account federal, state, local, and institutional grants. This seems the best way to help students and their families compare the financial aspect of college choice as well as to encourage colleges to focus on keeping costs down, especially for low-income students. Of course, admitting disadvantaged students and keeping costs low are only laudable if the institution helps students toward success after enrolling; and so the access and financial support institutions provide to disadvantaged students are best assessed in light of the outcomes these students experience. While it is important that colleges are affordable and provide access to disadvantaged students, it is essential that they contribute positively to the outcomes of their students.

Many commenters have made the point that a ratings system cannot capture all of the outcomes and values of postsecondary education. As discussed further below, there are many elements of education that are intangible or not amendable to a quantitative rating, or that are best determined and articulated by the institution or the education community.

That said, there are a set of outcomes widely embraced by students themselves and their families that provide a feasible starting point for improving the understanding of relative outcomes across institutions. Students want to know and compare institutions based on whether they will be able to complete their degree, find work, or go on to pursue graduate education; whether they will be able to repay the financial obligations they assumed to pursue their education; and whether they are likely to be able to earn a living and support their families.

Metrics being considered by the Department across all of the areas of access, affordability, and outcomes are discussed in more detail below. At this time some might be alternatives to each other, or we may decide to use related measures to generate a more complete picture of a domain than is possible with a single measure.

Percent Pell. The percentage of a college's enrolled students who receive federal Pell grants is the most common measure of access and the most readily available. Institutions annually report the share of their undergraduate students who receive Pell grants to IPEDS. This measure is simple, easily understood by families and the public, commonly evaluated by researchers and institutions, and amenable to longitudinal analysis because it has been collected for many years. Using Pell enrollment allows for a simple, clear measure of access for low-income students in particular, a core policy priority for the ratings system.

Pell enrollment data have some limitations. The fraction receiving Pell makes arbitrary distinctions between students from families with very similar family resources. In particular, for a given cost of attendance at an institution, there will be a cutoff level of expected family contribution (EFC)—a function of family assets and income—such that students just below the threshold qualify for Pell and those just above do not. This might give institutions an undesirable incentive to enroll students who are at the upper end of the income distribution among those who qualify for federal Pell grants and to avoid students from families just above the threshold who may not qualify for Pell. Research suggests there are growing gaps between low income and higher income students in college entry, persistence, and graduation rates.[1] As such, institutions may have an incentive to enroll low-income students who have the highest possible incomes while still qualifying as low-income.

EFC Gap. The EFC of a student is generally considered to be a strong measure of whether a student comes from a low socio-economic status or high socio-economic status background. It is also an important measure because it is used to determine the eligibility of undergraduates to receive Pell grants. Like Pell, EFC is calculated based on what families report on the FAFSA, including information about family income and assets. We are exploring defining EFC gap as the average difference between some focal EFC level and each student's individual EFC (with negative values treated as zero). Such a measure can be created from student-level data collected through the FAFSA and stored in NSLDS for federal student aid recipients only.

This measure would give institutions credit for enrolling students from lower family income backgrounds, but avoids creating sharply different incentives to enroll students from similar family circumstances

---

[11] "Inequality in Postsecondary Education" (with Susan Dynarski). In G.J. Duncan and R.J. Murnane (eds.), Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances. (Russell Sage: New York, New York, September 2011). Featured in the May 2012 NBER Digest (http://www.nber.org/digest/may12/may12.pdf).

simply because they fall just on either side of the eligibility line. The metric, however, has the limitation that it may misrepresent access for institutions with low participation in federal student aid programs. For instance, a two-year institution where few students receive Pell grants or federal loans (due to generous state grants and low tuition) may receive an access rating that understates its enrollment of low income students. Another concern is that it is more complex than the simple percent Pell measure.

Family Income Quintiles. We are also exploring family income quintiles as a possible measure of an institution's enrollment of low- to moderate-income students. Family income may be an important measure because it is the strongest factor in determining the EFC and financial aid eligibility of students. This calculation could also be derived from family income data reported on the FAFSA.

Like percent Pell, this metric would be simple and easy to understand, especially among consumers. Additionally, it may mesh well with a net price by income quintile metric, since both metrics include only federal student aid recipients and present information by income quintile. One downside of income quintiles is that, similar to EFC gap, family income information would be calculated based on what is reported on the FAFSA and may not accurately measure the presence of students in higher income quintiles because higher income students are less likely to receive federal grants or loans. Also, like percent Pell, a family income quintile metric may make arbitrary distinctions among students from families with very similar family resources near the borderline of each income quintile. This might again give institutions an undesirable incentive to enroll students who are at the upper end of the income distribution within each income quintile.

First-Generation College Status. While a metric based on income is the leading candidate for a basic access measure, we are also exploring other indicators such as the percentage of enrolled students who did not have a parent who attended college (i.e., first-generation college status). Such information is voluntarily reported by families on the FAFSA. Research has shown a link between parental education levels and child outcomes such as educational experience, attainment, and academic achievement. For example, there is a positive association between children with highly educated mothers and their rates of participation in early childhood education programs and home literacy activities.[2] In addition, children with more highly educated parents earn higher average reading and mathematics scores on the National Assessment of Educational Progress (NAEP) than do children with less-educated parents. [3]

Similar to percent Pell, this is a simple categorical measure that is easily understood by families and the public and commonly evaluated by researchers and institutions as an important factor for student outcomes such as completion. It also captures information on the extent to which institutions serve underrepresented populations of students that is not measured by metrics based on EFC or family income. However, parental educational attainment data have many of the same limitations as EFC gap. In particular, since the data available for this measure would be gleaned from the FAFSA, it would also be limited to federal student aid recipients, which could mean that the measure may be a less accurate assessment of the student body at institutions in which only a small fraction of students receive federal grants or loans.

Average Net Price. One measure being considered to ascertain a college's affordability is average net price, which provides information on how much students, on average, may need to pay to attend an institution each year. Colleges report average net price—the cost of attendance after accounting for all federal, state, and institutional grant aid—to IPEDS for first-time, full-time degree- or certificate-seeking

[2] http://nces.ed.gov/pubs2010/2010015/indicator1_5.asp
[3] Ibid.

undergraduates who were awarded grant or scholarship aid from the federal government, a state or local government, or an institution.

This measure comes closer to approximating the amount that students actually pay to attend an institution than "sticker price," but it is not complete. Students who are not first-time full-time or do not receive institutional grants, state grants, or any form of federal financial student aid are not included in the calculation. Such students include those who pay full price up front, rely on private loans, transfer in, or attend part-time. Furthermore, public institutions only report the average net price of in-state students, leaving out-of-state students out entirely. A more complete measure of college affordability would include average costs borne by all students who attend an institution, but such a measure is not currently available. Even though the IPEDS average net price measure falls short of an ideal measure, it provides meaningful information on the prices that many students actually pay to attend institutions.

Net Price by Quintile. Net price by quintile may offer a more accurate measure of the actual price paid by students of various family income backgrounds than average net price. Schools report to IPEDS on first-time, full-time degree- or certificate-seeking undergraduates who were awarded any federal student aid by quintile based on the student's family income. The income quintiles are broken down as $0-$30,000, $30,000-$48,000, $48,000-$75,000, $75,000-$110,000, and above $110,000. However, net price by quintile has many of the same limitations as average net price (e.g., currently limited to first-time, full-time and excludes out-of-state students). It also represents a more limited universe of students than average net price as it includes federal student aid recipients only, although this may therefore align better with outcomes and access metrics that are likewise limited to federal student aid recipients.

Average Loan Debt. At this point we are not considering using the average federal loan debt accrued by students while attending an institution as a measure of affordability. The Department recognizes the importance of such a measure and includes it on the College Scorecard to provide the information to consumers. However, there are several drawbacks to the measure that make it ill-suited for college ratings. First, student loan debt is partially a function of the financial resources available to the student. An institution serving significant numbers of low-income students, who need to rely on student loans to pay for college, could be rated as less affordable than another institution, which may have an even higher net price, because the second institution is serving higher income students who do not require loans to pay tuition. This could have a chilling effect on the willingness of institutions to enroll low-income students. Second, we are only able to measure federal student loan debt through NSLDS, so private loan debt could not be part of a debt-based affordability metric. This could create undesirable incentives for institutions to encourage students to substitute private loans for less costly federal loans.

Completion Rates. Completion is an important measure of college success, reflecting satisfaction of academic requirements, persistence, and capacity, and is associated with many positive lifetime outcomes from increased health and longevity to lower unemployment. Students who complete college also have higher earnings and are less likely to default on student loan debt than students who do not complete.

Colleges report completion rates for their fall cohort of first-time, full-time students to IPEDS on an annual basis. IPEDS completion rates are commonly cited and well-understood, but have well-known limits in that they fail to capture the range of outcomes relevant for many institutions and do not consider the completion outcomes of many students. For example, there are many institutions at which first-time, full-time students are a minority, sometimes a tiny minority, of students in attendance. Similarly, many institutions, particularly two-year institutions, have large numbers of students transfer to other institutions before graduating. While institutions that include transfer of students as part of their mission can report transfer rates to IPEDS, the ultimate completion outcomes of transfer out students are not reported in their graduation rates. However, this situation is temporary: starting with the current collection cycle, IPEDS will require institutions to report completion outcomes for part-time and transfer-in students and to distinguish between students who continue at another institution and those who fail to complete their education. The more comprehensive rates based on new reporting will not be publicly available until early 2017.

As an alternative to IPEDS based completion rates, we are considering developing completion rates using institution-reported administrative enrollment data in NSLDS that would capture outcomes among part-time students and transfers. Enrollment reporting has historically been used to determine when borrowers enter repayment, but can also be used to construct completion measures for students who receive federal student aid.

An NSLDS measure would allow us to incorporate completion outcomes for all federally aided students at an institution, including part-time students and transfer-in students. The data also allow us to construct completion measures for student subgroups, such as Pell students, based on data from the FAFSA. While the NSLDS completion measure may improve upon the IPEDS graduation rate, it is limited to only those students who received federal financial aid. The Department is currently conducting analysis to determine the accuracy of an NSLDS completion measure.

Transfer Rates. In addition to measuring completion rates, the Department is also exploring the viability of measuring transfer rates. While a four-year degree is, by definition, the highest undergraduate credential and an end in itself, a two-year degree is often not; for many students it is a step towards the next level of postsecondary education. To illustrate this point, consider two scenarios: in Scenario A, a student enters a two-year institution, receives a degree, and continues to a four-year institution; in Scenario B, a student enters a two-year institution, does not receive a degree, and transfers to a four-year institution. The first institution should undoubtedly receive credit for the student who completed and matriculated to a four-year program. The Department also believes that the second school should also receive credit for a positive outcome for the student transferring to the four-year institution, even without having completed at the two-year school.

In the case of a vertical transfer from a two-year to a four-year institution, we plan to reward two-year institutions for preparing students for transfer to a higher-level institution. This type of transfer is intrinsic to the community college mission and should be rewarded as such. Moreover, since many community colleges partner with four-year schools to allow for seamless transition from an associate's degree to a bachelor's degree program, we find no need to require the award of a credential so long as the student has gone forward to a four-year institution.

There are several important data limitations to note for the Department's ability to track transfer rates. Currently, institutions may voluntarily report transfer rates to IPEDS. The measure indicates the transfer-out rate of first-time, full-time degree or certificate-seeking students, which is the fraction of the total number of students who are known to have transferred out of the reporting institution within 150 percent of normal time (e.g., six years for four-year programs) divided by the number of students in the entering cohort. Alternatively, as with completion rates, the Department is exploring calculating transfer rates using student-level data from NSLDS. A current limitation with NSLDS is the lack of program-level data. As in the example described above, we would like to give credit to a two-year institution if the student transferred to a four-year institution, regardless of whether the student completed a degree or certificate at the two-year institution. Unfortunately, the Department will not yet know if the student transferred to a four-year program or to only a two-year program at the four-year institution.

While the Department is very interested in including transfer rates, there are several issues we are requesting additional guidance on if this metric were included in the ratings:

- *Should institutions receive credit for lateral transfers (i.e., from a two-year institution to another two-year institution) if the student goes on to graduate, if the student fails to later earn a degree, in both cases, or in neither case?*
- *Should institutions receive credit for transfers between four-year institutions? Would we say that preparation for transfer is a fundamental mission of four-year institutions and as such, we should treat lateral transfers as a positive outcome?*
- *Should the originating institution get credit for subsequent completions of transfer students (e.g., if a student transfers from a two-year to a four-year school, should the two-year get credit both for the transfer itself and also if the student were to complete their degree at the four-year)? Or should the two-year school only receive credit if the student ultimately completes a degree?*

Labor Market Success. An important goal of higher education institutions is to prepare students to enter the labor market successfully. Most students state that an important reason to go to college is "to be able to get a better job," and that getting a good job is an important factor in their college choice.[4] (Institutions with comparable student bodies and costs may have very different rates of employment and earnings, and part of the purpose of the ratings system is to help students as well as institutions and others appreciate those differences. At the same time, this is a complex topic: a better or good job can mean different things to different people. Are they able to secure work in their chosen field or location? Have they increased their earnings potential? Do they look for a myriad of other factors like responsibility, stability, a manageable schedule, status, or room for growth?

Many states have developed innovative tools to provide stakeholders with information about employment and earnings outcomes of their higher education institutions. These systems often rely on survey data or data from state unemployment insurance administrative data, both of which have limitations in terms of representativeness. The Federal government is uniquely able to provide nationally comparable information on employment and earnings for a broad swath of each institution's past enrollees that avoids some of the shortcomings of data from graduate surveys, or sources limited by state boundaries. However, the data are characterized by a different issue with representativeness, as such information is only available for students who have received federal student aid.

Some of the most extensive comments received about the ratings system have been on the subject of earnings measures. Many commenters noted the dangers of overemphasizing earnings for a variety of reasons. They expressed broad concerns about too great a focus on what people earn relative to other elements of work goals and rewards. They also discouraged approaches that would create excessive disincentives for students from entering fields that lead to important but lower paying positions, and noted the danger that schools might shrink their programs that prepare people for roles that are currently not highly paid, such as teaching and social and public service. Commenters also urged that the system not unduly reward schools in the measures for the decision of some of their students to go into highly lucrative positions. They cautioned that using measures of earnings soon after graduation may not be good measures of long-term capacity to earn a solid living. Commenters further noted that a student's major or program of study could influence future earnings. Labor market outcomes are likely to vary substantially within institutions (between program and majors), and because of this, it may be difficult to evaluate institutions fairly without having more complete information about the variance of earnings among programs within the same institution.

There were also comments flagging many variations that can affect the earnings of a school's graduates, including regional and rural-urban differences, along with patterns of discrimination against women, people of color, and others. Some commenters suggested using different metrics for ratings purposes,

---

[4] *The American Freshman: National Norms,* Higher Education Research Institute at UCLA, 2010, 2011, 2012

where a measure of earnings above some baseline might be appropriate, and for information purposes, where greater detail might be suitable for student use but not accountability.

To address these concerns, we are seeking ways to provide information about earnings in a way that gives a useful level of information for comparative purposes, but that avoids the detrimental effects noted by the commenters. In order to strike a balance among these measures, we are considering pairing a short-term indicator of "substantial employment" with a longer-term more specific earnings measure, such as mean or median earnings of former students ten years or more after entering the institution. Short-term labor market outcomes provide higher frequency feedback to stakeholders about institution performance, but long-term earnings outcomes more closely correlate with an individual's lifetime earnings and are thus a better proxy for career success.

A threshold measure of "substantial employment" would be a way of expressing the share of a school's graduates who earned above a specific level, essentially a "yes-no" question about earnings performance. For example, if the Department decided to use this "substantial employment" measure, we would likely use a measure such as the percentage of former students earning above 200 percent of the federal poverty line for a family of one. Another possibility for a threshold is a multiple of the full-time minimum wage earned over one year.

In either case, the Department would need to determine whether to measure earnings after a certain point from entering an institution, regardless of completion, or after a certain point from completing a degree. For example, among predominantly four-year institutions, we could look at short-term earnings six years after enrollment or two years after completion. For long-term earnings, in turn, we could look ten years after enrollment or six years after completion. In addition, in calculating long-term earnings there is an option to treat incomes above a certain high level as though they were at that level (i.e., topcoding), to reduce the influence of a small number of outliers such as students who become corporate executives after graduation. We would additionally exclude students who are enrolled in school in the earnings year, who are eligible for loan discharge due to total and permanent disability, or who have died.

Graduate School Attendance. In addition to completion and earnings, we are considering including a measure of graduate school attendance of former students within a period, such as ten years, from entry into the undergraduate institution. Colleges may have large numbers of students who go on to attend graduate school, itself an important outcome for an academic institution. While we are aware that the best source for this data is the National Student Clearinghouse, there is only one source available to the Department would indicate whether a student matriculated in graduate school: the presence of graduate loans in NSLDS. About 1.5 million graduate students took out Federal Stafford loans in the 2013-14 school year, representing roughly half of students enrolled in graduate programs. The Department is currently exploring whether it will be possible to include such a measure in light of the limited data.

Loan Performance Outcomes. The loan repayment or default behavior of former students may be included in the ratings system. Institutional cohort default rates are currently used as accountability measures for institutions of higher education, but other loan repayment metrics may provide additional or superior information on the ability of students to repay their federal loans. Options include adding rates of deferment and forbearance to default rates, measuring the percentage of students' initial loan balances that are repaid at various points in time, and measuring the percentage of students whose loans are in negative amortization at various points in time. These metrics can be calculated with NSLDS data.

Relatively simple metrics such as the repayment rate of borrowers (e.g., percentage of students reducing the principal on their loans within a window of time) can be important pieces of information as consumers weigh whether or not they will be able to handle their financial obligations after attending an institution. Loan repayment rates are not currently publicly available for most institutions of higher education, compounding the importance of providing these data to all students as they make one of the most important investments of their lives.

We believe loan repayment measures reflect student outcomes and institutional efforts to make college affordable in that they primarily reflect earnings, the cost of attending college, and student and family financial resources. They also provide some indication of whether institutions invest in student support services to help students manage the costs of their education. This may offer information about the well-being of graduates and the efforts of colleges to make college affordable, especially for students with fewer family resources to pay for their education.

There is some concern that loan repayment measures conflate outcomes and affordability with student and family financial resources, especially if earnings and the cost of attending college are separately factored into the ratings system. Accordingly, it may not be prudent to measure earnings and costs a second time under the guise of an additional outcome, particularly since loan debt is closely tied to student and family financial resources.

Performance over Multiple Years. In response to our RFI, many commenters expressed concerns about annual fluctuations in ratings due to anomalous one-time shocks or other sources of variability (e.g., missing data). To safeguard against volatility, several commenters suggested using rolling averages of metrics. Accordingly, we are exploring evaluating institutions based on performance over multiple consecutive years. We welcome comment on the best way to incorporate multiple years of data for each metric.

Intermediate Measures. The Department recognizes that most of the outcomes discussed above measure college performance with a lag of as much as six to ten years. The Department is therefore interested in comments regarding what intermediate measures, demonstrated by research to causally improve student outcomes rather than merely reflecting institutional wealth, might provide students and families with reliable and more timely information on institutional performance. These data could also help support and highlight institutions who are working to make continuous and measurable improvements that ultimately help move the needle on key student outcomes.

Outcomes Not Considered for Metrics At This Time. Many of the outcomes of postsecondary education are intangible, are not amenable to simple and readily comparable quantitative measures, and are not the subject of existing data sources that could be used across all institutions. Foremost among these are learning outcomes, which are central to understanding the value of an education but vary widely across programs and institutions and are communicated in many different ways. Other examples of information that may be of great interest to students and others but not easily captured include student and alumni public service and civic engagement; placement rates and patterns; licensure or certification passing rates; development of non-cognitive skills; and student satisfaction with the institution, generally and for specific services. For many reasons – including the desire for simplicity of the ratings system, institutional autonomy and differences, and lack of shared approaches and data – it seems preferable at this time to concentrate on the core data elements addressed above and to allow institutions to provide additional information to prospective students as discussed below.

How It Fits Together: In addition to receiving feedback on the individual metrics themselves, the Department is soliciting input on how metrics should be presented to measure performance and allow institutions to benchmark improvement, inform consumer choice, and, in the longer run, for purposes of accountability in the student financial aid programs. Rating an institution on each metric separately could make it easier to identify where an institution may excel and where it is relatively weak and may need improvement. As a result, keeping the metrics separate may allow for a more nuanced picture of varied institutions. For example, an institution that serves many disadvantaged students and has a high graduation rate might be considered as making important contributions to our system of higher education, even if it is costly to attend. On the other hand, without aggregating the metrics in some way, information about an institution's performance across multiple measures relative to the principles identified for ratings that comprise "value" may be difficult to assess for consumers and difficult to map in order to align incentives for an accountability purpose.

Key questions the Department is considering include:

- *Would consumers find it easier to see only a single rating for an institution, even if it meant less clarity about the varying levels of performance across metrics that may feed into the rating? To what extent could this be mitigated by providing both an overall rating and disaggregated metric information?*
- *A primary purpose of the ratings is to hold institutions accountable by measuring and publishing performance information. Would individual ratings on each metric or an overall rating make it easier to assess performance across the dimensions of the ratings system or would it hide or otherwise distort an institution's performance?*
- *Might it be useful to combine some related metrics (but not all of them) to allow for more a more meaningful representation of performance than possible on a single dimension in isolation? (For example, would it be helpful to merge into one element an outcome like completion rate, and the price students pay to attend the school?)*
- *What metrics based on combinations of the data described above would be most useful to guide consumer choice, and for institutional accountability?*

D. Adjusting Outcomes for Student and Institutional Characteristics.

While access and affordability are largely a direct result of college policies, the outcomes that students achieve may be influenced by the characteristics of the students themselves, like academic preparation and family resources; by institution-related factors, like quality of instruction and career services; and by other factors, like geographic location and the academic program the student chooses. In the outcomes domain, ratings would strive to take into account factors that are primarily unrelated to institutional policies and behavior. The influence of these factors on outcomes will, to some extent, already be mitigated by rating schools within institutional groupings, as discussed above. However, we are considering adjusting outcomes for important differences among institutions based on student and/or institutional characteristics to further diminish their influence on the ratings. The Department believes that ratings on outcomes that take into account the student bodies of institutions would be informative to prospective students and would be a fair way of assessing institutional quality. We also believe that taking student characteristics into account would help to avoid discouraging institutions from admitting and enrolling disadvantaged students. This concept, taken with a separate measure of access and recognizing institutions for improvement over time, could ultimately reward institutions that admit large numbers of these students.

Student characteristics. For college ratings, we are limited to extant federal data in attempting to reduce the influence of student characteristics on whether institutions are rated well or poorly on outcomes. For example, we have comprehensive individual-level administrative data on the income of students and their families, but we lack individual-level data on academic preparation (e.g., high school GPA or SAT/ACT scores). This means that we can take into account the socio-economic background of students who receive federal student aid within each college, but we cannot directly measure the academic preparation of individual students. We acknowledge that data limitations will impede our ability to fully take into account the characteristics of the students served by institutions, but the data we have will nonetheless enable us to provide a more fair assessment of institutional performance to the public than one that relies solely on raw outcome data.

Specifically, we are interested in developing a student-level regression model that would adjust outcomes (such as earnings and completion measures) for student characteristics that we can observe from the FAFSA. The student level data we are considering for this model include family income, EFC, independent/dependent status, parental educational attainment, age, has children, gender, marital status, veteran status, zip code/MSA characteristics, state of residence, transfer status, and enrollment intensity.

We strive to develop a fair way to recognize that institutions serve students with different backgrounds, while also developing a model that is simple and transparent as possible. We also would like to mitigate limitations of our data in order to ensure regression adjustments are robust despite lack of availability of key data elements, such as information about non-federal aid recipients and other direct measures of student preparation. With these considerations in mind, we welcome comments on several questions related to student characteristics:

- *In terms of the student-level data available to the Department, which student characteristics are the most important to include in a regression model to avoid harmful consequences for disadvantaged students and institutions that serve them?*
    - *Which student characteristics are likely to be most influential on outcomes?*
    - *Which student characteristics are likely to vary the most between institutions within the same institutional groupings? What kinds of student characteristics are unlikely to be adequately captured by institutional groups based on institutional characteristics such as predominant degree award level?*
- *Are some student characteristics more important to include for some outcomes metrics and some more important to include for others? Which student characteristics are important for all of the outcomes metrics we are considering, rather than just particular metrics?*
- *Given the limitations of the data we have, how should we assess the validity and reliability of the regression-adjusted outcome measures we are developing?*

Institutional characteristics. In addition to individual-level administrative data, we have institution-level data on student and institutional characteristics that could be used to further adjust outcomes. For example, admissions rate or average SAT/ACT scores could be used as an imperfect adjustment for academic preparation. Additionally, the majors and programs students enroll in and complete at an institution may affect the how much they earn after graduation. We are concerned that program mix may be correlated with earnings, even though the correlation between program mix and earnings may not accurately represent how well institutions are educating students. Thus, we are considering ways to adjust outcomes for institution-level characteristics, including measures of selectivity, the amount of funding public institutions receive from their states, institutional wealth (e.g. endowment size), and degree and program mix.

A potential risk of adjusting for institution-level data, however, is that to the extent that institutional policies and behaviors are correlated with institutional-level characteristics, adjusting for these factors may risk wrongly removing from the measures true differences in institutional performance. At the same time, we seek to mitigate the influence of student characteristics and other factors that we cannot take into account with student-level data. As discussed earlier, we are also interested in grouping similar institutions together based on many of the same institution-level characteristics as a way to supplement or complement the regression-adjustment of outcomes. For example, grouping similar institutions together and adjusting for a small number of important student-level characteristics may provide a model that is both transparent and fair, given the limitations of data available to the Department. With these considerations in mind, we welcome comments on the appropriateness of adjusting outcomes for institution-level characteristics.

E. College Ratings Website and Transparency Tools

The Department is considering ways to present the information about access, affordability, and outcomes in a clear way that increases transparency and meaningfully informs student choice. The Department seeks input from students and the public on the types of information, tools, and format that will meet this objective. Public input will inform the development of a ratings website and transparency tools that the Department plans to release by the 2015-2016 school year. Additionally, the Department will be doing consumer testing on the tools in order to ensure the ratings can be easily understood by prospective students and other users.

The Department recognizes that the broader community may benefit from, and possibly improve upon, the data and visuals made available through the college ratings system. As such, in addition to developing our own website and tools for the ratings, we are also investigating the best approach for developing an application programming interface (API) for the information included in the ratings system. This will allow other interested parties to lend a hand to our goal of providing meaningful information in an accessible and consumer friendly format through such tools as apps for smartphones.

Existing models of consumer-friendly ratings tools: The Department is aware that in order for the ratings system to succeed in informing consumers, institutional leaders, and policymakers, it is critical that the information be presented in a way that appeals to the end user. We also recognize that there are many other products and tools outside of the education sector that might offer guidance or good models regarding student and public information consumption. Finally, we are cognizant that this tool will not be released in a bubble, but rather into a field of existing educational tools, including those created by the Department such as the College Scorecard and College Navigator. The Department is particularly interested in comments on the questions listed below. We asked similar questions through the RFI and at the technical symposium but could benefit from additional guidance.

- *Are there examples of online tools that are particularly helpful for students and the public? What characteristics of these tools make them useful? How could the Department integrate similar tools and design features into the ratings website to make the information accessible?*
- *The Department publishes numerous existing tools already including the College Scorecard and College Navigator. Should the ratings system be incorporated into one of these existing tools or should the Department develop a new tool specific to the ratings? How should the Department publicize these tools to ensure that consumers are aware of them and their benefits?*
- *Is there specific research that would inform the effort to make our website and tools as user-friendly as possible?*
- *Are there examples from other fields that accomplish similar goals and could be good models for college ratings (e.g., energy efficiency "star" ratings, health provider ratings, nutrition labeling, etc.) ?*
- *How should the Department educate consumers on the use and meaning of the ratings?*
- *What kind of outreach campaign would be most effective in promoting the ratings?*

Consumer customization: The Department will be incorporating into the tool the ability for users to customize their own college ratings to reflect what is important to them. While the Department will be very clear about what metrics we feel are the most important for which institutions should be held accountable, we recognize that many consumers may want the opportunity to make their own value judgments about what is important to them. As such, the Department is requesting specific comments from consumers on the following questions:

- *Are there additional helpful categories by which to compare institutions? For example, would consumers want to be able to rate institutions by sector (e.g., public versus private), by location (e.g., within a specific state), etc.?*
- *Would consumers find it helpful to be able to select custom weighting for metrics? (That is, if the Department decides to roll-up the metrics into an overall rating, we may decide that one metric is more important than another one and weight it more heavily in determining an institution's rating but a consumer may want to apply equal weighting or even zero weighting to a metric.)*
- *Would consumers want to be able to see data on a measure disaggregated by a specific subgroup of students (e.g., Pell versus non-Pell recipients)? If so, what subgroups are helpful to include?*
- *How else might the customization tool be helpful?*

Additional information provided by institutions: Some commenters have suggested that, because so many aspects of postsecondary education and its outcomes cannot or should not be captured in a small set of metrics, the presentation of the ratings should provide an opportunity for colleges to add further information and links to their resources to provide a more complete institutional picture. While institutional websites and other materials often contain a great deal of additional information, it might be helpful to be able to refer or link to specific information directly from the ratings resource, potentially in conjunction with customized search results. This could take the form of an open ended space for additional information, a list of categories that could be populated by each institution (such as mission; learning goals and outcomes; public service measures; or breakdowns of ratings metrics by program, school, or major); or some other format. Some institutions might choose to use this opportunity to amplify measures in the ratings. For example, an institution might include four-year graduation rates if another metric is used for completion.

- *Would this approach be helpful to students and institutions?*
- *Would it be helpful to enumerate standard categories in addition to room for open-ended information? If so, what categories should be included?*

III. Next Steps

Thoughtful and comprehensive feedback is important to designing an effective ratings system. The Department welcomes comments until February 17, 2015 on the approaches and specific questions outlined in this document either publicly through the comment section of the Department's higher education blog or by email to collegefeedback@ed.gov. Comments will be reviewed and used to inform discussion and analysis throughout the next several months, so early submissions are especially encouraged.

The Department will also initiate and participate in structured discussions to help illuminate specific issues and to understand the effects of options under consideration for particular purposes, users, populations, and types of institutions. Sessions that will be open to the public will be announced in early January through the Department's website.

The Department will use public and expert responses on the framework and metrics to develop or revise the metrics and system. At that stage it will go on to incorporate data and generate institutional ratings, leading to the system that will be available in advance of the 2015-16 school year.