CLASP Comments on the College Ratings System Draft Framework
Submitted to the Department of Education on February 17, 2015

The Center for Law and Social Policy (CLASP) is pleased to offer comments on the Department of Education's draft framework for a college ratings system. CLASP—a national nonprofit that develops and advocates for policies that improve the lives of low-income people and works to increase the number of low-income adults and youth who earn postsecondary credentials—has made a number of recommendations related to the ratings system over the last year. These include written comments, testimony, a briefing paper on implementing a system that empowers students while avoiding unintended consequences, and a briefing paper on the importance of presenting workforce outcomes. We are pleased to see some of our recommendations addressed in the draft framework.

First and foremost, we commend the careful attention paid to the potential unintended consequences of a ratings system on low-income and underprepared students. While the framework does not have all the answers on this account, it is asking the right questions. Second, we are gratified that the Department plans to include measures of labor market success among its metrics. Students, especially low-income students, go to school to improve their earnings potential – so a postsecondary education ratings system without workforce outcomes would be sorely insufficient. Last, we appreciate the attention to creating fair comparison groups of institutions that take into account differences in institutional characteristics and missions. The strategy of grouping colleges and universities by predominantly two- and four-year institutions is a good start, and the framework rightly identifies additional characteristics for consideration like program mix and admissions selectivity. Determining how to take institutional and student characteristics into account using peer groups will be a challenging but critically important task.

We are concerned, however, that the framework fails to respond to the call from CLASP and many other organizations to make a distinction between uses of data for transparency as opposed to accountability. It is very important to distinguish the data elements and presentation most appropriate for each purpose. While considering transparency separately from accountability may potentially more time consuming, a single system that is intended to serve multiple purposes may either increase the chances of unintended consequences for low-income and underprepared students, or provide little useful information to students, and possibly both. For these reasons we recommend that the primary purpose of the ratings system should be to provide consumer information.

In the following comments, we address selected issues in the framework in more detail, in the order in which they are found in the framework, focusing on the topics of: 1) institutional groupings, 2) labor market success measures, 3) separate or aggregate metrics, 4) adjusting outcomes for student and institutional characteristics, and 5) consumer customization. Finally, we offer recommendations for two additions to the draft framework: program-of-study level data and voluntary data reporting.

Institutional Groupings

We support the framework's intent to create peer comparison groups so that only broadly similar institutions are being compared. This is important to reduce disincentives to serve low-income and underprepared students. At a minimum, the plan to group predominantly two-year institutions separately from predominantly four-year institutions, due to difference in missions, is a good start. As the Department considers accounting for additional characteristics, CLASP recommends grouping institutions along key institutional differences that have strong predictive power for the outcome metrics. These dimensions are discussed in detail below in our comments on Adjusting Outcomes for Students and Institutional Characteristics. Using robust comparison groups is better than adjusting outcomes for student outcomes and institutional characteristics. The concepts are highly related, and the Department should think about them together, as we pair them below.

Labor Market Success

CLASP is very pleased to see measures of labor market success included in the draft framework. Workforce results matter to students, who consistently report that their prospects for employment and improved earnings are very important to them as they enter college.[1] According to a newly released survey from the Higher Education Research Institute at UCLA, three of the top four reasons that 2014 freshman cited as "very important" in deciding to go to college are related to labor market success: getting a job (86.1%), training for a career (77.1%), and making more money (72.8%).[2]

While including measures of labor market success is important, great care must be taken to avoid unintended consequences related to access for low-income and under-prepared students. If measures are too stringent, then institutions may become more selective, leaving behind those in most need.

From a cursory scan of nine state websites that measure postsecondary education students' employment and earnings outcomes, we noted 14 unique measures of labor market success, including full-time employment, employment paying at least minimum wage, mean and median first year earnings, and earnings two, three, five, and ten years out of school, among others. While mean first year earnings is the most commonly used measure across the sites, the breadth of measures shows that there is clearly no commonly held measure of choice.

Any discussion of the "best" earnings or employment measures will depend on the overarching policy objectives, as well as the specific context within which each measure will be used. For instance, uses of workforce outcomes measures may include:
- Consumer information
- Performance measurement for continuous improvement
- Performance measurement for accountability
- Performance-based funding

It appears that the Department is trying to find measures that can meet multiple policy objectives used in multiple contexts. CLASP supports separating measures used for transparency

from those used for accountability. However, if measures must meet the purposes of both transparency and accountability, to minimize unintended consequences for low-income and under-prepared students, certain considerations should be taken:

- Earnings levels for accountability purposes should not be set so high as to discourage the use of career pathway programs that may have lower earnings levels for initial credentials, but the potential of higher earnings later (e.g., nursing assistant).
- Earnings results should be presented in a labor market context if possible. For example, median earnings results could be supplemented with information on the percent of a state or regional median wage that the earnings result represents.
- Earnings results should always be presented for particular programs of study, rather than for an entire institution (see Consumer Customization, below).

Employment measure

"Substantial employment" as described in the draft framework would indicate whether or not former students achieve a certain level of earnings, for example 200 percent of the poverty level for a family of one. This construct makes sense for a metric to support accountability, since it could be used to identify a minimal level of acceptable annual earnings. Using such a level is better than counting an individual as employed if they have any earnings in the time period under consideration, which is how the workforce development performance measures generally work, because it would set too low a bar.

Earnings measure

The longer-term earnings metric described in the draft framework would measure mean or median earnings of former students ten or more years after entering the institution. This could be an appropriate metric for consumer information purposes, but we believe ten years is too long. Five years would be a more reasonable amount of time for prospective students to use in decision making. If this metric were used for accountability purposes, while ten years might better reflect the lifetime earnings of a former student, five years may correlate just as well.

Measurement timing

The draft framework states that labor market success metrics would have to be measured either some number of years after entrance or some smaller number of years after graduation. We understand that the data on graduation rates is often not timely or accurate, but using NSLDS program-level data may be feasible as the graduation data improves in this database. A third option would be to measure some number of years after exit, meaning graduation or leaving school without a credential. Many students upgrade their skills and obtain employment well before obtaining a credential, and a graduation-based earnings metric may exclude these students. For instance, the new common indicators of performance under the Workforce Innovation and Opportunity Act (WIOA) are measured a number of quarters after exit. Such an approach would work only if good data were available for when individuals exited school.

The most feasible option may be to measure some number of years after entrance, which, while it has pitfalls, may be the only option with reliable data. If the earnings metric is based on a time period following enrollment, rather than graduation, it should include different time periods for different program durations, in order to capture post-program earnings that are more comparable.

Separate or Aggregated Metrics

In addition to asking for input on metrics themselves, the Department has also asked for comments on how metrics should be presented, and, specifically, if it would be helpful for consumers or policymakers to have a single overall metric that takes into account all of the individual metrics. CLASP opposes the creation of a single composite rating. Because not all metrics are equally important to different consumers, individuals should be able to view the individual metrics most important to them. The federal government should focus on data, metrics, and groupings, and on making data available to outside entities with expertise in providing consumer information to students.

From an accountability perspective, a composite rating may be difficult to understand and difficult to use for program or institutional improvement. Treating institutions fairly would require that any disqualification thresholds be based on clear criteria. These criteria should be designed to remove those institutions that fail to meet minimal standards of performance over time. Setting thresholds based on a single criterion can create strong incentives to game results or distort institutional missions and may also have other undesired results. No institution should ever be confused about the minimal threshold or metric(s) they must improve to retain Title IV eligibility.

Adjusting Outcomes for Student and Institutional Characteristics

Consistent with our recommendation that the primary focus of the ratings system should be on consumer information, rather than accountability, we do not recommend adjusting actual outcomes to account for differences in the types of students who are enrolled or for institutional characteristics. For consumer information purposes, we think that only unadjusted information should be provided to students and stakeholders. Consumers should always be able to see what the actual institutional result was for each of the metrics used in the rating system. However, consumers should be able to use selection criteria to compare institutions based on location and other factors, as we have discussed above, and they should be able to compare the results for each metric with the results of other peer institutions. These peer groupings should be comprised of institutions that are similar in terms of the types of students enrolled, the institutional mission, and other characteristics that strongly affect their results. In our view, the most appropriate function for statistical analysis in support of the rating system will be the formation of these peer groups.

Selecting the student and institutional characteristics that will be used in forming peer groupings will be a very challenging task. There are student characteristics that may be highly relevant (e.g., level or preparation for college work) but for which data are not generally available. There are characteristics that may be essential for disaggregating results (e.g.,

race/ethnicity, gender) but that should probably not be used to form comparison groups, to avoid setting lower expectations for student results based on membership in one of these groups. Finally, there are institutional characteristics that are highly relevant for determining results (e.g., level of selectivity in admissions), but that may be difficult to measure
.

   Given all of these considerations, CLASP recommends the following student and institutional characteristics be explored for creating peer groupings for comparison of results:

   Student characteristics:
   - Pell recipient, or recipient of other need-based aid;
   - Low-income based on FAFSA data;
   - First-generation college status; and
   - For earnings results, program of study.

   Institutional characteristics:
   - Level of selectivity (e.g. percent of applicants admitted);
   - Primary types of credentials granted (awards, certificates; associates, bachelors, advanced); and
   - Percent of students attending other than full-time.

   If the system will be used for accountability beyond setting minimum thresholds of performance for certain metrics that all institutions are expected to meet, then it must have some process for setting institutional expectations that takes into account the differences in critical institutional and student characteristics. Without incorporating such protections for institutions that enroll low-income students and help them succeed, an accountability system will create perverse incentives to enroll and focus resources on the most prepared students and those most likely to succeed in postsecondary education and the job market – thus harming the most disadvantaged and underprepared students. Until the Department can determine how to take institutional and student characteristics into account using peer groups, we strongly recommend that the ratings system not be used for accountability purposes.

   Regression-based adjustment models have been used for years by the U.S. Department of Labor (USDOL) in setting <u>expected</u> outcomes for workforce programs (JTPA, WIA, now WIOA), but the context for these programs is very different than what pertains to a college ratings system. USDOL sets expected outcomes that are meant to reflect a realistic goal for actual performance, as well as promoting improvement over time, rather than a minimal threshold intended to eliminate the poorest performers. Furthermore, the consequences for failing these goals are significant, especially for local program administering entities, so there is tremendous pressure to meet whatever expectation is set. This is almost certainly not the context for the ratings system in the foreseeable future. We urge the Department to continue its dialogue with the community, including workforce programs at the federal and state levels, to further its understanding of the most appropriate uses of these statistical methods for setting expectations.

Consumer Customization

According to the draft framework, the Department will be incorporating into the ratings tool the ability for users to customize their own college ratings to reflect what is important to them. Consumers would want to see data on a measure disaggregated by specific subgroups. This is in keeping with CLASP's recommendation that students be able to view results for students like themselves. We recommend the following subgroups:

- Program of study: This would be helpful for each of the metrics, but is of particular important for labor market success measures, which are most useful when presented in the context of a program of study.
- Pell Grant recipients: Does the institution obtain good results for both Pell recipients and non-Pell recipients?
- Full-time/part-time/mixed status: Students' ability to attend full time heavily affects their prospects for graduation. It is a disservice to students not to make this reality clear. Further, students should be aware that some schools are more successful than others with part-time students.
- Gender: Showing institutional results for students by gender, particularly in settings where they will be underrepresented, is an important part of the context.
- Race/ethnicity: The ability of the institution to minimize the achievement gaps for minority students is an important element for comparison.

Suggested additions to the framework: Program of Study-Level Data and Voluntary Data Submission

While CLASP appreciates the incorporation of some elements of our past recommendations in the draft framework, we were disappointed that the Department has not included disaggregation of data by program of study. Program of study-level data are useful for all of the metrics, but are especially important for measures of labor market success. Prospective students should be able to see how students in various programs of study fared in the labor market, to help inform their choice of program. Non-traditional and low-income students, in particular, need data by program of study, because they are often place-bound due to family or work constraints. They are unlikely to choose to leave home to pursue their postsecondary education; however, they can choose among the many options of programs of study at a local university, community college, or other conveniently-located institution.

We understand that the National Student Loan Data System (NSLDS) has recently begun collecting students' program of study, but that the data may not be ready to use because of quality control. CLASP recommends that the Department use the program of study data it has as soon as practical. If a ratings system started to use the NSLDS program of study data, it would create a strong incentive for institutions to improve the quality of this information.

A second recommended addition is to allow institutions to voluntarily report data on students who are not required to be in the NSLDS. As students who did not qualify for aid, these students are likely to be better prepared and will probably have higher outcomes than students

currently in NSLDS. If these students have better outcome measures, it would be in institutions' best interest to voluntarily report the data. NSLDS may not be currently configured to allow such voluntary reporting, but movement in that direction could address the concern about only having data on students who receive federal financial aid.

Thank you for the opportunity to provide comments. If you have any questions, please contact Anna Cielinski at 202-906-8019 or acielinski@clasp.org.

---

[1] Tim Harmon, Neil Ridley, and Rachel Zinn, *Workforce Results Matter: The Critical Role of Employment Outcome Data in Improving Transparency of Postsecondary Education and Training*, Center for Law and Social Policy, 2014, http://www.clasp.org/resources-and-publications/files/2014-04-29-CLASP-Workforce-Results-Paper.pdf

[2] Kevin Eagan, Ellen Bara Stolzenberg, Joseph J. Ramirez, et al., *The American Freshman: National Norms Fall 2013*, Higher Education Research Institute at UCLA, 2015, http://www.heri.ucla.edu/monographs/TheAmericanFreshman2014.pdf