Using National Data in University Rankings and Comparisons

Denise S. Gater

TheCenter Reports (June 2003)
http://thecenter.ufl.edu

Using National Data in University Rankings and Comparisons

Denise S. Gater

This is a summary covering two presentations on the above topic. The first was presented by Denise S. Gater at the annual meeting of the AAU Data Exchange in Columbus, Ohio (May 2, 2003), and the second was presented at the Association of Institutional Research (AIR) annual Forum in Tampa, Florida (May 20, 2003) by Denise S. Gater and Diane D. Craig.

Introduction and Background

Various university constituencies generate an ever-increasing demand for comparative data to support state-mandated accountability efforts, performance funding initiatives, accreditation reviews, academic program reviews, national benchmarking efforts, as well as the controversial university rankings published by commercial enterprises such as *U.S. News & World Report*.

This paper shares the results of our experience in working with the national data and highlights some of the limitations of these data when used for comparative purposes. This is especially critical when various agencies and institutions attempt to use these data for the purpose of measuring and comparing university performance. The information we present here draws primarily on *TheCente*r's work on *The Top American Research Universities* project, work by the University of Florida's Office of Institutional Research, and the experience of institutional researchers at several other research universities.

TheCenter staff have focused particular attention on the university financial data from the IPEDS Finance survey (collected annually by the National Center for Education Statistics, Washington, DC). Last year, we examined university financial reports and compared them with IPEDS Finance survey data for a specific group of research universities because university resources are so critical to a research university's ability to be competitive (for more details, see *The Top American Research Universities*, August 2002). Our study demonstrated the unreliablility of much of these data and found institutional comparisons based on these unreliable data highly misleading.

During 2003, *TheCenter* staff collaborated with institutional researchers from four other major research universities (University of Arizona, University of Kansas, University of Illinois-Urbana Champaign, and Purdue University) to share our experience in attempting to use data from the IPEDS Finance survey for comparative purposes. We shared highlights of those discussions with the Association of American Universities Data Exchange (AAUDE) at its meeting on May 2, 2003. In that presentation we provided a review of the issues associated with using these data to an audience frequently asked to construct, explain, or debunk the various financial ratios based on IPEDS Finance information routinely published by universities and other higher education groups. Due to time constraints of the

presentation, we focused on public universities using the "pre-GASB 34/35" IPEDS Finance survey form.

In a broader presentation to the Association of Institutional Research (AIR) annual Forum in Tampa (2003), *TheCenter* staff discussed key challenges and problems in using national data for university comparisons and incorporated examples from IPEDS Finance data and faculty counts from the IPEDS Salaries survey. In addition, we touched on a variety of methodological issues and gave an overview of how we have used national data on universities for *The Top American Research Universities* project.

Key Challenges and Problems

Some of the basic challenges that researchers face when using national university data for comparisons include:
1) measuring at the university level versus measuring at the program level;
2) comparing characteristics that are not comparable; and
3) ensuring appropriate measures for the intended audience.

Of particular importance, data that measure an institution's performance at the university level (such as total research expenditures) is fundamentally different from data that measure the productivity of the faculty of that institution (such as average faculty research productivity). The first takes the institution as the unit measured where the second takes the faculty as the unit measured. Given the differences in size and scope, composition, and mission, comparisons of faculty performance at the university level present an even greater challenge than conducting comparisons at the program level.

Program/discipline level data sometimes provide effective measures of faculty productivity. For example, accrediting agencies, such as the American Chemical Society, routinely provide national data on faculty productivity that are useable for comparisons, such as the average number of refereed publications per chemistry faculty member. They calculate these ratios quite precisely and chemistry departments generally accept these measures as reliable for national comparisons. However, for purposes of comparing universities, normalizing data relative to a faculty size almost always produces highly misleading results since reliable faculty counts do not exist nationally. We will illustrate the limitations of these national faculty data later in this paper.

Institutional comparisons also suffer from the difficulty of selecting comparable peer institutions. The complex and diverse nature of universities requires careful analysis to ensure that institutional comparisons match reasonably similar universities. Institutional size and scope, mission, mix of disciplines, institutional control (public/private), presence or absence of a medical school, agricultural extension and/or agricultural experiment station, and data for a single campus versus multiple campuses or systems all have a profound effect on the comparability of universities.

Unfortunately, the IPEDS Institutional Characteristics (IC) survey data do not easily distinguish the presence of a medical school, and studies that rely on these data can produce erroneous conclusions. For example, an institution that awards first-professional degrees, a common selector for medical education, may not have a medical school because first-professional degrees also include degrees in dentistry,

law, pharmacy, veterinary medicine, and others. The IPEDS Completions survey is more reliable in this case because it identifies institutions that award medical degrees.

Nationally available data often contain information on single campus institutions, multiple campus institutions, and state university systems, but without clearly identifying the universe the data includes. Comparisons using these data can be unreliable without adjustments because a comparison of a single campus institution with a multi-campus institution can occur. In order to increase the validity and usefulness of these data, we strongly recommend adjusting the raw data when necessary in order to reflect information for equivalent units (for example, *TheCenter* attempts to present data only for single campuses by adjusting system level data to reflect the individual campus components.)

Finally, comparative studies of universities must consider the intended audience when selecting institutions and data elements for analysis. Two primary user groups for this type of information are:

> 1) prospective students and parents (targeted by college guides and commercial rankings, such as *U.S. News & World Report*) and
> 2) university administrators and institutional researchers who want to examine institutional performance.

Publishers of commercial rankings, of course, sell their publications to college-bound students and parents. To enhance the sense of currency of their publications they may introduce new measures or change the weighting of measures from one year to the next. They also typically incorporate unreliable data based on reputation surveys in their ranking profiles.

While comparisons and rankings published by the media may be of interest and may serve their intended audience, they do not usually provide data that allow universities to measure their relative performance over time. Data can be a powerful management tool for evaluating and improving university performance, not only for comparing an institution against itself from year to year but also for comparing an institution to its competitors over time. The commercial surveys do not generally serve this purpose.

Data and Methodological Issues

Reliability is the foremost concern when using national data for comparing university performances. Unreliable data can appear due to differences in institutional accounting practices, limitations on reporting ability, and variations in institutional interpretation of data definitions. As a result, these data may not measure what the users intended to measure. In particular, data from two national sources prove especially difficult for university comparisons: financial data collected in the IPEDS Finance survey; and faculty counts from the IPEDS Salaries survey.

National data used for comparisons must be clean, timely, consistent, and stable. Most federal agencies and national organizations now collect data via the web, and their timeliness in making these data available to the public has improved for the most part. Even though the organizations have processes in place to check and clean the institutional data they collect, researchers must also verify the data. By examining an institution's data over time, inconsistencies appear quickly and

researchers can follow up regarding issues of suspected inaccuracies or reasons for apparent anomalies. In addition, further analysis may allow the replacement of missing data with estimates or substitute figures from another source.

The decision of whether to display a list of rank-ordered institutions versus displaying groupings of institutions is a key methodological issue. Commercial rankings, in particular, favor providing a rank-ordered listing of colleges and universities because their purpose is to create the illusion of significant differences among closely ranked institutions and in particular to focus attention on small changes in rank order. However, by assigning a precise order to institutions, these listings incorrectly imply that exact differences actually exist from one institution to the other. Such fine distinctions do not exist in reality and *TheCenter*'s comparative analysis, for example, focuses on groups of institutions reflecting performance on a number of significant measures.

University quality and performance require multiple measures, but determining a set of reliable indicators using nationally available data presents a challenge. Many performance measurement systems prove unsatisfactory because they collect and report too much information. A large number of often highly intercorrelated measures makes the analysis more difficult to interpret and the project difficult to manage. In *The Top American Research Universities* project, *TheCenter* limits the number of measures to nine carefully selected, key measures.

Problems with Using IPEDS Finance Data for Comparisons: Universities, higher education groups, the media, and others often attempt to use ratios of "revenue per FTE" or "expenditures per FTE" as performance measures by extracting revenue and expenditure data from the IPEDS Finance survey, since this is currently the only national source of financial data available to researchers. An example of such a calculation from Florida's Council for Education, Policy, Research, and Improvement appears in the figure below.

$$\frac{\text{1999-2000 Revenue}}{\text{1999-2000 FTE}}$$

Revenue =
State Appropriations + Tuition & Fees
+ Local Appropriations
(Source: IPEDS Finance, 1999-2000)

Source: "Equity in Funding in the State University System of Florida," by the Council for Education Policy, Research, and Improvement (CEPRI), 2002.

The numerator in the calculation above comes from the numbers reported to IPEDS Finance, and matches figures reported in the university's annual financial statement. While the numbers accurately reflect the revenue, the composition of the revenue differs dramatically by institution, rendering the ratios useless for comparative purposes. For example, applying the above calculation to the University of Florida does not reveal that 35 percent of the numerator comes from revenues for the Agricultural Experiment Station and Agricultural Extension (Institute of Food and

Agricultural Sciences) and the Health Science Center, two units most of whose revenues relate to the administration and support of these units, not revenues for instruction. Moreover, most other universities will not have these revenue items, even if they have the same number of students, rendering any comparisons invalid.

$652 million = Total Numerator for UF

- Less $117 million for Institute of Food & Agricultural Sciences
- Less $113 million for Health Science Center

Resulting ratio changes from $15,856 to $10,276

If we correct this calculation by removing the unrelated revenues for these units from the total revenue figure for the University of Florida as illustrated above, the ratio changes dramatically and places the university in an entirely different comparative group among other institutions within the state or nationally.

As mentioned earlier, *TheCenter* recently reviewed IPEDS Finance data and annual financial reports for a group of 11 public universities and 11 private universities and found numerous problems with the reliability of these data for university comparisons. In this discussion, we focus on the public institutions. *TheCenter* review included the following private and public institutions.

Private	**Public**
California Institute of Technology	Pennsylvania State University
Columbia University	University of California - Berkeley
Cornell University	University of California - Los Angeles
Duke University	University of Florida
Harvard University	University of Illinois - Illinois-Champaign
MIT	University of Michigan - Ann Arbor
Northwestern University	University of Minnesota - Twin Cities
University of Chicago	University of North Carolina - Chapel Hill
Stanford University	University of Virginia
Yale University	University of Washington - Seattle
University of Pennsylvania	University of Wisconsin - Madison

Many observers assume that for public universities, state appropriation data represent comparable data and mean the same thing at all public institutions. On the contrary, it turns out that the figures reported as appropriations may or may not include main campus only, branch campuses, medical centers, and agricultural extension/agricultural experiment station appropriations. Unfortunately, the IPEDS Finance report does not indicate what it includes in the state appropriations category for each institution.

IPEDS Finance also does not clarify where the following revenues appear and probably the institutions report them in different ways: appropriations derived from auxiliaries; agency funds; appropriations for capital outlays; and investment income. As an example, although university auxiliaries are supposed to be self-sufficient, we found wide discrepancies between auxiliary revenues and auxiliary expenditures for universities, ranging from a $32 million surplus to a $53 million deficit. Public university foundations may or may not be included in the revenue section due to differences in accounting practices. Indeed, an institution may indicate zero endowment income in this section of the report if, for example, the foundation is an independent "component unit" of the university. Again, the data may be correct for accounting purposes but not comparable with other institutional reports and therefore inappropriate for comparative purposes.

The IPEDS survey uses a vague definition of the revenue categories "Sales and Services of Educational Activities" and "Other Sources of Revenue" providing no clear method for determining what, exactly, universities report in these categories and inhibiting useful comparisons. Probably, IPEDS should combine these two categories since many universities appear to use them interchangeably.

The expenditure portion of the IPEDS Finance survey presents similar problems for those trying to use the data to compare institutions. The data do not accurately reflect institutional spending in the categories of research and instruction because fund accounting rules result in either exaggerating or underreporting the true instructional or research costs. This major weakness comes from the practice of assigning 100% of a fund to either teaching or research if 51% or more falls into one or the other category. Institutional comparisons based on these data will almost certainly distort the institutions' performance.

Varying financial conventions and policies among states also lead to incomparable data among public universities. In some states, expenditures for such things as debt service, worker's compensation, fringe benefits, retirement, and utilities may fall to the university and in others the same expenditures may appear in the state's accounts and not the university's. In particular, we found that the figures reported for debt service in the IPEDS Finance survey often do not match the annual financial statements of institutions. Again, we cannot determine how much of the debt service the state carries independently of the university. This represents a major difference in institutional financial reporting because capital expenses represent a significant issue for all higher education institutions. Other expenditures that may or may not appear in a university's IPEDS Finance data include the costs of a central system office, intercollegiate athletics, and the expenses of a fund-raising foundation. Even if included in a university's data, IPEDS does not indicate which expenditures appear and which do not. All of this makes comparisons based on IPEDS financial data particularly unreliable.

Research expenditures may not accurately reflect an institution's true expenditures on separately budgeted research because of various accounting practices, and research expenditures for work conducted by affiliated or university-owned hospitals may or may not get counted in the university's data. We compared research expenditures reported in the IPEDS Finance survey versus figures reported on the NSF R&D Expenditures survey and found large differences for some institutions. Yet, IPEDS and NSF both state that schools should report separately budgeted research expenditures. The following table compares data from these two surveys for a subset of research universities. Although the NSF data only report science and engineering disciplines, this does not explain the disparity between the two sets of data. For example, the difference at University at Buffalo appears to result from unusual system reporting practices. The SUNY system office completes the IPEDS Finance surveys and includes only research expenditures that processed through its research foundation, even though the campus does not process all sponsored research through the research foundation. The campus completes the NSF R&D Expenditures Survey and captures and reports all these expenditures. As a result, the University at Buffalo's sponsored research is dramatically under-reported in the IPEDS Finance data.

Comparison of IPEDS Finance and NSF Research & Development Expenditures (Selected National Universities-Doctoral)

Institution	1999 NSF Total Research (x $1,000)	1999 IPEDS Total Research (x $1,000)	IPEDS $ / NSF $
Massachusetts Institute of Technology	420,306	621,080	148%
Stanford University	426,549	539,861	127%
Harvard University	326,193	397,183	122%
New York University	167,179	203,213	122%
Carnegie Mellon University	142,174	143,310	101%
Princeton University	124,237	125,028	101%
Georgetown University	111,426	109,939	99%
Emory University	189,170	182,493	96%
Vanderbilt University	149,675	140,938	94%
Columbia University	279,587	262,869	94%
Tulane University	87,324	79,522	91%
California Institute of Technology	212,216	187,234	88%
University of Wisconsin - Madison	499,688	440,483	88%
University of Pennsylvania	383,569	334,072	87%
University of Virginia	157,487	136,058	86%
University of Texas - Austin	258,122	221,142	86%
Duke University	348,274	296,856	85%
Rice University	41,069	34,632	84%
Cornell University	395,552	328,727	83%
University of Washington - Seattle	482,659	400,332	83%
University of Chicago	162,805	132,565	81%
University of California - Santa Barbara	104,561	84,061	80%
Northwestern University	233,809	182,171	78%
Case Western Reserve University	182,332	141,111	77%
University of Michigan - Ann Arbor	508,619	388,898	76%
Brandeis University	48,305	36,551	76%
Yale University	274,050	205,476	75%
University of California - San Diego	461,632	345,919	75%
University of California - Irvine	141,842	105,385	74%
University of California - Los Angeles	477,620	351,942	74%
University of Illinois - Urbana-Champaign	358,247	262,907	73%
Washington University	315,606	229,911	73%
University of Rochester	177,126	127,277	72%
University of Southern California	280,741	200,711	71%
University of California - Davis	307,950	215,688	70%
University of North Carolina - Chapel Hill	252,767	174,973	69%
University of California - Berkeley	451,539	289,253	64%
Texas A&M University	402,203	247,256	61%
Brown University	76,330	40,808	53%
Tufts University	100,872	42,055	42%
SUNY-Buffalo	166,823	61,581	37%

Sources:

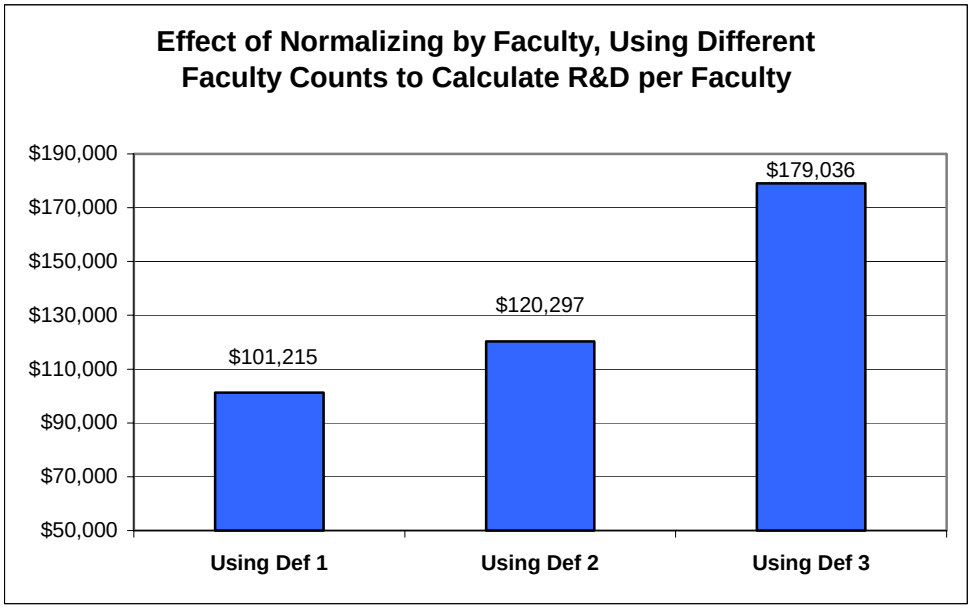
NSF/SRS Survey of R&D Expenditures at Universities and Colleges, FY99.
NCES IPEDS Finance Survey, FY99.

The University of Kansas Office of Institutional Research has done extensive work over a 25-year period on peer comparative data, including visits to their peer institutions, to learn how they are organized and how their data are reported. They used IPEDS Finance data as one component of their peer selection process and found striking differences in how different institutions report their financial data. Based on these experiences, they strongly urge researchers to draw on case studies to better understand financial data, and they also suggest that campus institutional researchers meet with financial staff members who prepare IPEDS Finance data at their own institutions to better understand data reporting conventions and make recommendations for any changes in reporting.

We have focused here on the "pre-GASB 34/35" reporting format from NCES. NCES developed and adopted new IPEDS Finance survey form for public institutions to accommodate the rule requiring that all state and local government agencies implement the new GASB 34/35 accounting standards by FY2004. The form is optional now and will be mandatory in Spring 2005. The new form is substantially different from the previous one. For more information on the specific changes, see [http://nces.edu.gov/ipeds/web2000/GASB.asp]. Even with the change in reporting, we expect that many of the same comparability issues will remain.

Problems with Using IPEDS Salaries Data for Comparisons: We mentioned earlier the importance of differentiating data that measure an institution's performance at the university level from data that measure the productivity of the faculty of that institution. When we compare universities by the amount of federal research expenditures, for example, we provide a market share indicator because we are measure how much of the total federal research belongs to that institution's productivity. However, when a study normalizes data by the number of faculty, they provide a reference to the productivity of individuals within the institution. A university may control a large portion of the total federal research expenditures, but because it has a large faculty who do many things besides research, it may have a low productivity per faculty member. The methodology that normalizes university productivity by faculty numbers and then compares university performance on that basis distorts the data because the denominator of faculty numbers does not represent the same universe relative to research in all universities. This methodological flaw would invalidate these comparisons by itself, but in addition the data on number of faculty are also highly unreliable.

The most commonly used data source for the faculty counts is the IPEDS Salaries survey. The primary problem with using faculty counts from IPEDS Salaries is that the methodologies used to produce these counts vary significantly from one institution to another, making the use of these data for comparisons misleading and unreliable. This survey intends to collect salary data only on the subset of "instructional faculty" in an institution. Differing interpretations of the IPEDS definitions for "instructional faculty" can have a major impact on any resulting ratios constructed with these faculty counts. And using such counts to normalize research productivity measures produces poor analysis in almost every case.



Source: University of Florida, Office of Institutional Research.

The example above shows the effect of using different interpretations of definitions to report what is supposedly the same population of faculty *in one institution* – the number of full-time instructional faculty. We should point out that faculty counts of "instructional faculty" are inappropriate to use in devising a ratio of R&D per faculty at the university level and the above is for illustration purposes only. If one did want to look at comparative research productivity among institutions, we would need a standard measure to identify the full-time-equivalent research faculty across institutions.

We used the following definitions to construct the ratios in the above table. According to our informal inquiries to research universities across the nation, some research university used each of these definitions to report faculty in their IPEDS Salaries survey.

- Def 1 - all full-time ranked faculty (professor, associate, and assistant), excluding only clinical medicine
- Def 2 - all full-time tenure/tenure-track faculty
- Def 3 - ranked faculty who teach 50% or more, excluding all medicine

A paper by *TheCenter* staff titled "The Use of IPEDS/AAUP Faculty Data in Institutional Peer Comparisons" provides more details on this topic and is available on *TheCenter* website [http://thecenter.ufl.edu/gaterFaculty1.pdf].

Conclusions

Universities need to use the apparently comprehensive data provided by federal and other agencies with great care. While institutional research offices in every university make great efforts to report their data accurately and completely, the conventions used, the differences in context in each university, the different accounting systems, and the sometimes inconsistent or vague guidelines provided by various agencies make these data extremely difficult to use effectively. The complexities and errors identified in this paper provide useful cautionary examples to help those who use these data recognize the weaknesses in the underlying information. While the temptation to use such large-scale databases is strong, great caution in making significant generalizations or comparisons based on these data should become the constant watchword of all institutional research. This paper offers some general guidelines and principles for good practice when using nationally available university data to measure and improve university performance.

What we have shared here is a product of the data and collective experience of institutional researchers at various universities as well as by *TheCenter* staff. *TheCenter*'s work relies on the insight and recommendations of many colleagues throughout the country who contribute data, information, and perspective. The effectiveness of these techniques has brought national attention and a commitment to translate the methodology from particular implementations at various universities to a general data-driven perspective applicable to any research university. For more information on *TheCenter's* work and to access *TheCenter* data and publications, visit [http://thecenter.ufl.edu/].