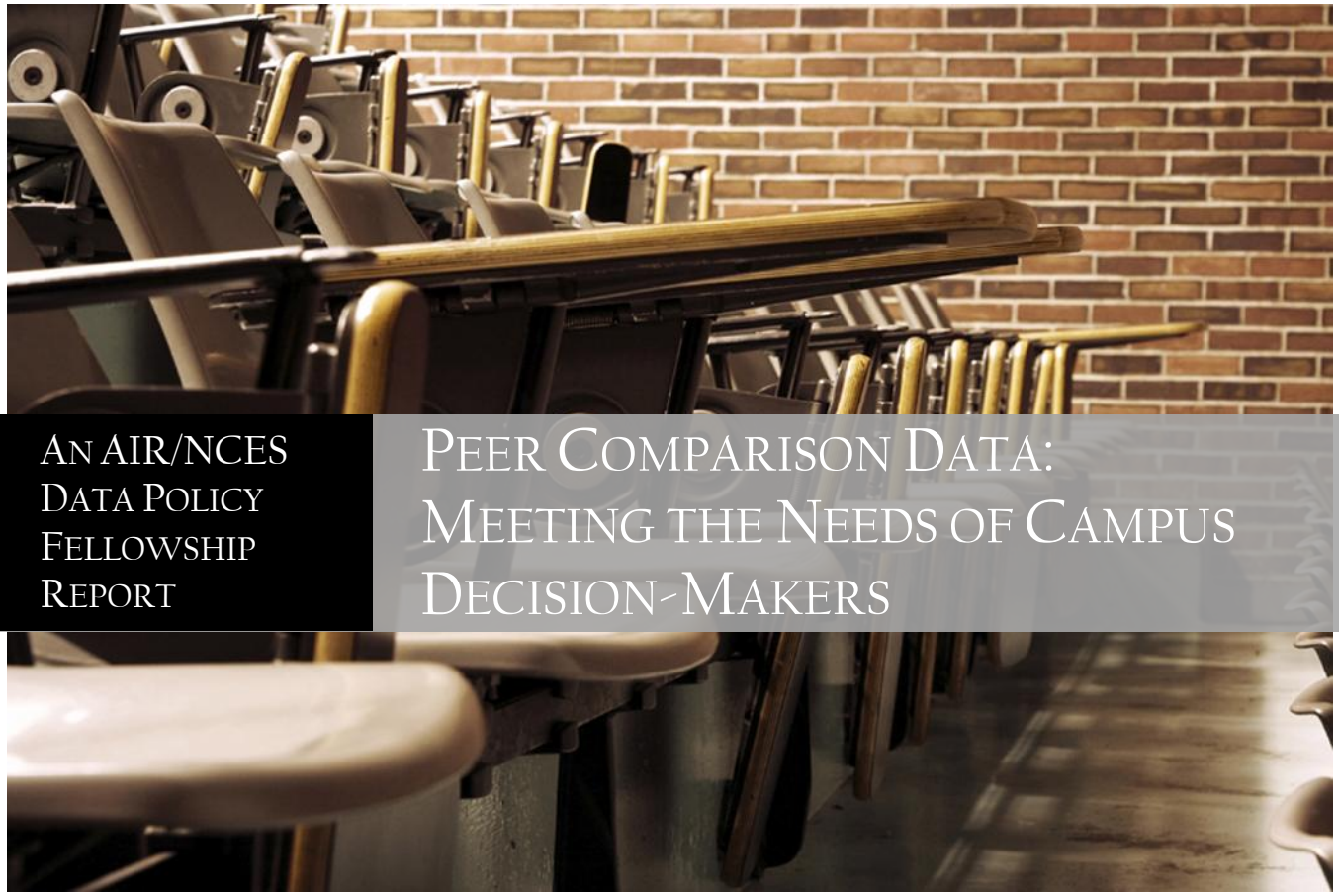
Peer Comparison Data: Meeting the Needs of Campus Decision Makers

Matthew Soldner

AIR/NCES Postdoctoral Policy Fellowship

July 2009

This report was prepared for the National Center for Education Statistics at the US Department of Education under Contract Number ED-03-CO-0020.



An AIR/NCES Data Policy Fellowship Report

Peer Comparison Data: Meeting the Needs of Campus Decision-Makers

Matthew Soldner | July 2009

The AIR/NCES Policy Fellowship Program

The National Center for Education Statistics (NCES) Postsecondary Studies Division collects and analyzes data on postsecondary education in the United States. As part of fulfilling its charge to "collect, analyze, and report education information and statistics in a manner that…is objective…and… is relevant and useful to practitioners, researchers, policymakers, and the public" (Public Law 107-279 § 151), NCES funded the Policy Fellowship Program through the Association for Institutional Research. Policy fellows plan and conduct a year-long research project designed to result in improvements to the quality, comparability, and usefulness of the Integrated Postsecondary Education Data System (IPEDS).

Contents

Why Focus on Peer Comparison Data?

Whether it takes the form of graduation rates, indicators of student learning, or evidence of financial efficiency and effectiveness, information about what is occurring on college campuses is in high demand. Increasingly, aggregate information about institutional performance is used not only to paint a descriptive portrait of a single campus, but also to facilitate comparisons of that performance relative to a set of policy benchmarks or the mean (or median) performance of a set of peers.

Three decades of critique levied at American public education, combined with resource scarcity, has driven today's emphasis on the use of postsecondary education data for evaluative and comparative purposes. Much of the former has come about due to national dialogs spurred by so-called "reform reports," demanding increased diffusion of data on educational quality to stakeholders (transparency), who in turn can use that data to made a wide range of policy judgments (accountability) (n.b., The Association of American Colleges and Universities, 2002, 2007; The Secretary of Education's Commission on Excellence in Education, 1983; The Secretary of Education's Commission on the Future of Higher Education, 2006; The Wingspread Group on Higher Education, 1993). When the policy-maker is the consumer of educational services, the relevant judgment may be the choice of a particular college. Indeed, the Department of Education's CollegeNavigator, which provides consumers at least some the information they need to form a rational judgment (e.g., net costs of attendance, degrees awarded, and graduation rates) about the utility of attending one college versus another, represents a prototypical response to transparency-accountability pressures.

Similarly, in the face of real fiscal constraints, some state higher education authorities have further impelled the transparency and accountability movement by comparing institution-level data with pre-established benchmarks for the purpose of adjusting campus budget allocations, a practice known as performance funding (Burke & Minassians, 2002; Burke & Serban, 1998). While this practice is not yet widespread (Carey & Aldeman, 2008), private sector support for such schemes continues (e.g., Lumina's "Making Opportunity Affordable" program). Flat or worsening economic conditions would also seem to make performance funding—or other systems of coupling state appropriations to college outcomes—an increasingly attractive policy position. Notably, performance funding already exists in Florida, Ohio, Pennsylvania, Texas, and Virginia, as well as a handful of other states.

Finally, the process of accreditation can, through the self-study process, encourage institutional benchmarking. Middle States (2007), for example, advises institutions to engage in peer comparisons along a number of metrics, including student learning, financial resource use and distribution, and faculty recruitment and retention. Similar guidance is offered by WASC (2008) in their *Handbook of Accreditation* and by SACS (2004) in their *Handbook for Reaffirmation of Accreditation*. Programmatic accreditors may mandate some level of peer comparison, or simply

encourage it. The former is exemplified by the Association to Advance Collegiate Schools of Business (2007), who actually prepare a peer report as part of their reaccreditation process, while the latter is the approach suggested by the American Psychological Association's Committee on Accreditation (2008).

Taken together, these trends suggest that the demand for peer comparison data will, at the very least, remain healthy. Given recent legislative history, one could readily speculate that demand will increase. And where there is a demand, supply is sure to follow.

The Role of National Center for Education Statistics

Charged with the responsibility to "collect, analyze, and report education information and statistics in a manner that … is objective … and … is relevant and useful to practitioners, researchers, policymakers, and the public" (Public Law 107-279, Section 151), the Department of Education's National Center for Education Statistics (NCES) is a major provider of data surrounding the status of postsecondary education in the United States.

Particularly relevant to those interested in peer comparison is NCES's Integrated Postsecondary Education Data System (IPEDS). A repository of data about all Title IV-participating colleges and universities in the United States, IPEDS annually collects information about institutions' characteristics, enrollments, completions, staffing levels, finances, and costs of attendance. Because it gathers institution-level data on a variety of metrics, is mandatory for the majority of traditional colleges and universities, is relatively timely and is readily available to anyone with access to the Internet, IPEDS data in its raw form is a potentially rich source of peer comparison data.

Since 2004, IPEDS data has been diffused via a product known as the *Data Feedback Report* (DFR). Produced yearly for each institution in the IPEDS universe, the DFR is a print product that contrasts the IPEDS data submitted by an institution with its institutionally-defined comparison group. In the absence of such a group, a "generic" group developed by NCES on the basis of any number of institutional characteristics (i.e., 2005 Basic Carnegie Classification, sector, region, urbanicity, degree-granting status, and, for program-reporters, largest instructional program) is substituted. The comparative metrics presented on the DFR are derived from all survey components, although the number of comparisons included in each report varies by institutional sector and size. In mid-October, links to each institution's DFR are e-mailed to its IPEDS keyholder. At the end of October, paper copies of the DFRs are mailed to campus chief executives. A related product, the Executive Peer Tool (ExPT), allows users to print duplicate DFRs, change the composition of comparison groups and re-generate DFR figures and tables, or access supplemental figures not included in the default report.

Types of Alternative Providers of Peer Comparison Data

NCES is only one source of peer comparison data available to campus decision-makers. At least four other broad classes of providers exist in today's marketplace:

1. Single-state data systems. Although necessarily limited in scope to the state in which they are housed and, often, to the specific institutional control of the sponsoring agency (e.g., public colleges supported by a state agency or independent colleges supported by an advocacy organization), single-state data systems can provide a robust source of comparative data. The scope of data contained in these systems can vary widely, both in terms of granularity (i.e., student unit record versus aggregate information for an entire institution) and metrics. Notable publically available systems include the California Community Colleges Chancellor's Office's Data Mart[1] and the State Council of Higher Education for Virginia's Data Warehouse.[2]

2. Narrowly focused, academically-oriented research. Several university-sponsored research programs engage a sufficient number of campuses across the country that they provide, at least within the universe of participating institutions, a mechanism for benchmarking. Typically, these programs are focused on a narrow range of metrics (often some element of the student experience) that are directly tied to the research interests of the faculty members associated with them. Examples include the Higher Education Research Institute's Cooperative Institutional Research Program (CIRP[3]), sponsored by UCLA, or the National Survey of Student Engagement (NSSE[4]), sponsored by the University of Indiana.

3. Organizational data sharing activities and comparison products. Numerous organizations, including professional associations, scholarly societies, advocacy groups, and disciplinary accreditors, offer members access to one or more activities or products that, directly or indirectly, facilitate institutional comparison. While these activities and products may emerge as a hallmark feature of an organization (e.g., NACUBO's Endowment Study) rarely do they dominate the organization's mission. Examples of full-fledged benchmarking include the Association of Governing Board's Benchmarking Service[5] or the Council of Independent College's Key Indicators or Financial Indicators Toolkits[6]. Instances of data sharing conducted for purposes other than benchmarking but that nonetheless facilitate inter-institutional comparisons include the American Society for Engineering Education's On-Line Engineering College Profiles[7].

---

[1] See http://www.cccco.edu/CommunityColleges/DataMart/tabid/848/Default.aspx, retrieved July 10, 2009.
[2] See http://research.schev.edu/, retrieved July 10, 2009.
[3] See http://www.heri.ucla.edu/, retrieved July 10, 2009.
[4] See http://nsse.iub.edu/index.cfm, retrieved July 10, 2009.
[5] See http://www.agb.org/wmspage.cfm?parm1=1113, retrieved July 10, 2009.
[6] See http://www.cic.edu/projects_services/index.asp, retrieved July 10, 2009.
[7] See http://www.asee.org/publications/profiles/index.cfm, retrieved July 10, 2009.

4. Single-purpose data consortia/exchanges. Unlike data sharing projects sponsored by organizations with a broader charge, single-purpose data consortia/exchanges exist for the sole purpose of facilitating inter-institutional data sharing. Historically, consortia/exchanges have emerged as institutions coalesce around a common characteristic or concern. Examples include the Association of American Universities' (AAU) Data Exchange (AAUDE), which is limited to AAU members, the Council on Financing Higher Education (COFHE), which is limited to a small group of elite, private institutions, and the Southern Universities Group (SUG), whose membership includes large, public institutions in the south. The types of data shared among consortia members can vary widely, and may include: (a) early-release IPEDS data, (b) data collected as part of another survey (e.g., the NACUBO Endowment Study), and (c) original data collections developed by consortia members.

Questions for Research

As the federal statistical agency charged with collecting, analyzing, and diffusing policy- and research-relevant information concerning the status of postsecondary education in the United States, NCES should understand how its IPEDS data products—and in particular the DFR—are used by campus-based practitioners, and how those products might be improved. The availability of alternative sources of postsecondary education data suggests both that practitioners turn to non-NCES products to meet their data needs and that, from their own experiences, alternative providers may have developed "good practices" from which NCES can learn.

The purpose of this research was three-fold: (a) to understand how campus senior executives use the DFR, (b) to understand the extent to which campuses make use of alternative providers of peer comparison data, and (c) to better understand the scope and activities of alternative providers of peer comparison data. To fulfill that purpose, the following research questions were posed:

1. What is the role of IPEDS data in meeting campus executives and IR officers' needs, including:
    a. administrators' ratings of the importance of currently available IPEDS metrics,
    b. the sources to which administrators' turn to access IPEDS data,
    c. with whom IPEDS data are shared (e.g., legislators, Regents, boards of visitors, campus leadership teams, faculty and staff, students, the media), and
    d. the ends to which administrators use IPEDS data?
2. How are institutions using alternative providers of peer comparison data consortia, to include:
    a. whether they belong to one or more consortia,

   b. their use of data shared within consortia (e.g., student; finance and management; human resource; facilities and assets), and
   c. the perceived "value added" of consortia membership?
3. What does the universe of peer alternative providers of peer comparison data "look like," to include:
   a. a description of providers' extent, membership, and purposes, and
   b. a portrait of services offered (e.g., data collection and analysis, training and professional development, ad-hoc evaluation), data products provided (e.g., reanalysis of IPEDS data or peer comparison group development), and focal topics addressed (e.g., student; finance and management; human resource; facilities and assets)?

Brief Methodology

To address this project's three research questions, a suite of three studies were developed and implemented between July, 2008 and July, 2009. They included:

1. The Campus Executives' Study (CES). The CES was a web-based survey of 1334 senior executives at four- and two-year, degree granting institutions across the United States. Almost 500 (n = 491, 37% response rate) executives responded to a series of questions focused on their use of the DFR and their judgments about the utility of DFR metrics.

2. The Keyholder Supplement (KHS). The KHS was a web-based survey of 817 potential[8] IPEDS keyholders at four- and two-year, degree granting institutions across the United States. Almost 400 (n = 394, 49% response rate) invited participants responded, along with 220 "convenience" responders solicited through e-mails sent to This Week in IPEDS and the Association of Institutional Research (AIR) E-AIR newsletter. Questions on the KHS paralleled those found on the CES, and also included a series of specific questions related to institutional use of alternative data providers.

3. The Data Providers Study (DPS). The DPS was a web-based survey of 375 organizations, associations, and consortia serving postsecondary education. More than 200 (n = 206, 55% response rate) invited participants responded, indicating whether their organization offered comparative services and, if so, providing information on those services' scope, data sources, and distribution methods.

Each study was supplemented by a qualitative component. After completing the survey instrument itself, respondents were asked whether they were interested in participating in a

[8] Because a list of keyholders was not available from NCES, respondents were identified by targeting individuals with job titles commonly associated with involvement in IPEDS data collections, including those in institutional research, registrars, and budget and planning. Among invited participants, 76% reported they were their institution's keyholder. For more information, see Appendix A: Detailed Methodology.

follow-up interview. Approximately 10 interviews were conducted after each study. When appropriate, interview data is summarized in this report to augment quantitative results.

Findings: The Campus Executives' Study

The goals established for CES, in order of precedence, included assessing the extent to which campus executives: (a) recognized the DFR, (b) shared it on and off campus, (c) perceived the broad classes of metrics included in the DFR as important to their work, (d) turned to various sources to access important metrics, and (e) used the DFR in their completion of their job responsibilities.

Institutional characteristics of respondents. Half (50%) of respondents represented public institutions, with the balance consisting of respondents from private, not-for-profit (37%) and private, for-profit institutions (8%). More than half of respondents (58%) represented 4-year and above institutions, with 37% representing two-year institutions. Slightly more than 5% of respondents' level and/or control were unknown, due to refusal or logical error (e.g., a respondent indicated working at a less-than-2-year institution, even though that sector was not sampled). Greater rates of participation among public institutions, and lesser rates of participation among private, for-profit-institutions, were observed than statistically expected if response had been independent of sector, $\chi^2(2)=68.505$, $p \leq .001$.

Respondents reported a median unduplicated headcount of 3200 students, with the majority of respondents reporting headcounts between 1142 (the 25th percentile) and 10007 (the 75th percentile) students. This suggests that respondents come from institutions that are generally larger than the sampling frame from which they were drawn (25th percentile: 686 students, median: 1958 students, 75th percentile: 6183 students).

Position of respondents. Although invitations to participate in the CES were sent to campus chief executives, respondents were allowed to redirect their invitation to another senior administrator on campus save one: because the CES and KHS covered similar topics and were simultaneously fielded, recipients were specifically asked to *not* delegate the completion of the CES to institutional research officers. Ultimately, the overwhelming majority (69%) of CES respondents identified themselves as their campus's chief executive. Other job titles reported by respondents included: (a) provost or chief academic affairs officer (9%), (b) institutional research officer (8%), (c) other senior executive (4%), (d) chief financial officer (3%), (e) chief of staff (2%), and (d) enrollment management officer (2%). Post-hoc tests suggested that, in the public, 4-year sector, delegation was observed more frequently than statistically expected, while the opposite was true in the public, 2-year sector, $\chi^2(5)=20.129$, $p \leq .001$.

Receipt and review of the DFR. To determine whether questions pertaining to their use of the DFR were relevant, after providing demographic data respondents were shown an image of the 2008 DFR's cover-page and asked whether they recalled receiving it, either directly from NCES

or from a colleague on campus. Eighty-six percent of respondents indicated that they had received the DFR. (Those respondents who did not report having received the DFR skipped to questions related to their ratings of the "Importance of Institution-Level Metrics," described below.)

After confirming receipt of the DFR, respondents were asked whether they reviewed the report. Virtually all respondents (91%) reported that they had reviewed the DFR. Those who reported not reviewing the DFR cited several reasons, including: (a) delegation (n = 18), (b) determining it was not a priority (n = 8), (c) a lack of time to do so (n = 6), and (d) a perception that it was not relevant to their institution (n = 3).

An important caveat.
Caution should be used when drawing inferences from the rate of review reported above. To the extent that an institution's decision to upload a customized comparison group is an indicator of the importance placed on the Data Feedback Report by institutional users, the rate of customization—although rising—is inconsistent with this study's findings.

At the associate's level and above, approximately 26% of institutions customized their comparison group in 2007-2008. Rates vary by sector, with four-year publics reporting the highest rate of customization (~52% in 2007-2008). Institutional control is also an important determinant of customization: at the associates' level and above, only around 5% of for-profit institutions customize their DFR comparison group in 2007-2008.

Diffusion of the DFR on campus. Respondents who reported having received the DFR were asked about whether they shared it with selected parties. Table 1, below, details chief executives' experiences with the diffusion of the Data Feedback Report.

A review of the table below suggests that campus executives diffuse their institution's DFR within a fairly narrow range. Approximately 80% of chief executives share the DFR with members of their cabinet, other direct reports, and institutional research staff, although rarely (between 15% and 25% of the time) is the president actively leading a discussion on its content. DFR data tends not to be shared beyond the immediate administrative ranks, however. Between 60% and 94% of chief executives reported no use of the DFR with regents, campus governance (e.g., faculty, staff, or student councils), campus media, state or local officials, or accreditors.

Why the DFR is appropriate for some groups and not others is unclear from the data. Although not typically indicated, some degree of speculation may be appropriate here. For example, it may be that chief executives are concerned that, without additional context, those not intimately familiar with the institution or IPEDS might tend to draw inappropriate inferences about the DFR's content. Whether that *is* the case is, in some regard, irrelevant: Providing additional context to data products is almost *always* beneficial to their users. Other plausible concerns could include the derivation of individual metrics, the creation of comparison groups, or the magnitude of particular differences. Irrespective of what the concern truly is, if we have reason to suspect any or all of them might be operating to minimize the DFR's utility, then developing approaches to address them seems appropriate.

Table 1. Campus Executives' Experience with Data Feedback Report Diffusion

	Respondent reported having …			
Venue, Sorted in Order of Presentation	No Use	Forwarded	Joined a review	Led a review
President's Cabinet	18%	41%	16%	25%
Direct Reports	22%	42%	15%	22%
Institutional Research Staff	17%	45%	23%	15%
Regents/Visitors	66%	15%	7%	12%
Campus Governance	63%	19%	10%	8%
Campus Media	72%	18%	5%	5%
State/Local Officials	94%	3%	1%	1%
Accreditors	93%	3%	3%	1%

Notes.
Totals may not sum to 100% due to rounding error.

Importance of institution-level metrics. Irrespective of their use of the Data Feedback Report, all respondents were asked to rate the importance of a series of broad metrics, including those available through the Data Feedback Report. Three non-IPEDS-derived metrics, including workforce development, technology use, and faculty productivity, were included based upon Burke & Minassian's (2002) work on the data needs of institutional research officers and state higher education officials.

Table 2, below, details respondents' ratings, sorted in order of the frequency with which respondents rated the importance of each metric.

Table 2. Campus Executives' Ratings of Institution-Level Metrics

Metric, Sorted in Order of Presentation	Important	Somewhat Important	Somewhat Unimportant	Unimportant
Student persistence and attrition	82%	13%	3%	3%
Student recruitment and admissions	70%	22%	6%	3%
Student cost of attendance and aid	65%	26%	6%	3%
Faculty salaries and benefits	64%	24%	7%	5%
Core revenues and expenses	63%	26%	8%	3%
Faculty workload and productivity[a]	55%	33%	8%	4%
Faculty diversity	54%	31%	10%	6%
Graduate job placement and workforce devel.[a]	48%	34%	12%	7%
Development and endowment	45%	34%	12%	9%
Instructional technology[a]	42%	44%	11%	4%
Physical plant and capital assets	37%	42%	16%	6%
Space utilization[a]	33%	39%	22%	6%

Notes.

[a] Data on this metric is not available through an IPEDS survey component.

Given that importance ratings may suggest future directions for NCES data products, detecting statistically significant differences between pairs of ratings is helpful. Although a clear trend of decreasing importance is evidenced as one moves down the table, naively detecting statistically significant differences in the ratings of pairs of metrics is difficult (e.g., whether there is a "real" difference in chief executives' ratings of student persistence and attrition versus their ratings of student recruitment and admissions). In Table 2a, below, results of McNemar's comparison of two dependent proportions highlights statistically significant differences between pairs of metrics. More darkly shaded cells indicate greater levels of statistically significant difference and answer the question: "Is the difference between these two percentages statistically significant in this sample?"

As can be seen in Table 2a, three metrics evidenced particularly high chi-square values, albeit for very different reasons: student persistence was rated comparatively high, while physical plant and capital assets and space utilization were rated comparatively low. Table 2a also demonstrates a lack of discrimination between the importance of various faculty metrics.

Table 2a. Statistically Significant Differences Between Executives' Ratings of Institution-Level Metrics

	Chi-Square Test Statistic For McNemar Comparison of Two Dependent Proportions											
Comparison	Faculty Salaries	Faculty Diversity	Faculty Work	Student Recruit	Student Persist.	Student Costs	Rev. & Expense	Physical Plant	Space Utilization	Devel. & Endow.	Instruct. Tech.	Work. Devel.
Faculty Salaries												
Faculty Diversity	19											
Faculty Work	17	6										
Student Recruit	8	28	28									
Student Persist.	42	92	96	41								
Student Costs	6	22	18	7	48							
Rev. & Expense	9	16	11	11	59	4						
Physical Plant	88	45	57	129	197	107	130					
Space Utilization	101	62	81	143	214	123	140	15				
Devel. & Endow.	48	17	28	76	144	69	65	26	34			
Instruct. Tech.	54	30	27	86	166	59	64	14	39	24		
Work. Devel.	30	5	18	69	143	46	41	19	42	5	21	

Notes.

[a] McNemar-Bowker test is computed with six degrees of freedom.

$\chi^2 \geq 32$	Significant at $p \leq .001$, with Bonferonni adjustment
$32 > \chi^2 \geq 27$	Significant at $p \leq .01$, with Bonferonni adjustment
$27 > \chi^2 \geq 23$	Significant at $p \leq .05$, with Bonferonni adjustment
$23 > \chi^2$	Not statistically significant with Bonferroni Adjustment

Sources of data. Individual respondents who rated any metric as at least "somewhat important" were then asked to indicate to which of the following sources they turned to gather the data: (a) the Data Feedback Report, (b) the NCES DataCenter website, (c) a consortia of peer institutions, or (d) an off-campus association or company. Responses are summarized in Table 3, below, sorted in order of frequency with which the Data Feedback Report was cited as a source of data for a given metric. Metrics marked with an "[a]" are generally not available through any IPEDS survey component.

Two features are of relevance for NCES. First, on those metrics addressed by IPEDS, the Data Feedback Report was "competitive" with most off-campus data sources (faculty salaries are a notable exception), although the gap between the DFR and respondents' first choice option was great (between 40% and 70%). Second, NCES Data Center Tools were, in every instance, respondents' fifth choice. Whether this is a function of utility (perceived or actual) or marketing is not clear, at least on the basis of CES survey data.

Table 3. Campus Executives' Use of Various Data Sources

	Percent of Respondents Reporting Use of a Given Data Source			
Metric, Sorted by use of DFR	DFR	NCES DataCenter	Consortia of Institutions	Off-Campus Association or Company
Faculty Salaries	28%	14%	40%	29%
Student Persistence & Attrition	28%	18%	27%	17%
Student Costs & Financial Aid	28%	19%	29%	20%
Faculty Diversity	27%	14%	27%	17%
Student Recruitment & Admissions	24%	14%	26%	19%
Core Revenues & Expenses	22%	11%	21%	15%
Faculty Workload [a]	17%	9%	21%	14%
Development & Endowment	16%	7%	22%	23%
Job Placement & Workforce Devel. [a]	16%	7%	18%	21%
Physical Plant & Capital Assets	15%	6%	16%	11%
Instructional Technology [a]	15%	7%	20%	18%
Space Utilization [a]	13%	4%	14%	12%

Notes.

[a] Data on this metric is not available through an IPEDS survey component.

Uses of institution-level data. Next, executives were asked how (if at all) institution-level data were used in their "day-to-day practice." Respondents were able to select any or all of the following four options, each believed to represent important, interconnected university business processes: (a) resource allocation, (b) strategic planning, (c) peer comparison, and (d) stakeholder communication. Table 4, below, summarizes their responses, sorted in order of the metrics use business processes involving peer comparison.

Although respondents' use of various data types for the purpose of peer comparison is necessarily due, in part, to the availability of that data, Table 4 suggests something of the metrics on which peer comparisons are most frequently done (e.g., faculty salaries, student costs, student persistence and attrition, and faculty diversity).

Table 4. Campus Executives' Use of Institution-Level Data in Key University Business Processes

	Percent Reporting Use in Business Process			
Metric, in Order of Use for Peer Comparison	Resource Allocation	Strategic Planning	Peer Comparison	Stakeholder Communication
Faculty Salaries	72%	71%	69%	45%
Student Costs & Financial Aid	59%	78%	64%	64%
Student Persistence & Attrition	60%	84%	61%	60%
Faculty Diversity	32%	75%	60%	53%
Student Recruitment & Admissions	67%	81%	57%	58%
Faculty Workload	64%	69%	51%	30%
Job Placement & Workforce Devel.	44%	74%	48%	66%
Development & Endowment	62%	80%	47%	62%
Core Revenues & Expenses	56%	78%	44%	47%
Instructional Technology	71%	78%	36%	38%
Physical Plant & Capital Assets	71%	76%	33%	44%
Space Utilization	65%	79%	27%	31%

To identify statistically significant differences between respondents' use of various metrics for the purpose of peer comparisons, Table 4b, below, summarizes all pair-wise comparisons of the proportions reported above. More darkly shaded cells indicate greater levels of statistically significant difference and answer the question: "Is the difference between these two percentages statistically significant in this sample?"

As can be seen there, the reported rate of use of a metric like "Faculty Salaries" for the purpose of peer comparison—approximately 69%—was statistically significantly different than the rates for all other metrics with the exception of faculty diversity, student persistence, and student costs.

Table 4b. Statistically Significant Differences Between Executives' Use of Metrics for Peer Comparison

	Chi-Square Test Statistic For McNemar Comparison of Two Dependent Proportions											
Comparison	Faculty Salaries	Faculty Diversity	Faculty Work	Student Recruit	Student Persist.	Student Costs	Rev. & Expense	Physical Plant	Space Utilization	Devel. & Endow.	Instruct. Tech.	Work. Devel.
Faculty Salaries												
Faculty Diversity	11											
Faculty Work	38	10										
Student Recruit	18	1	4									
Student Persist.	8	1	16	5								
Student Costs	3	2	17	10	3							
Rev. & Expense	65	23	6	22	45	49						
Physical Plant	96	59	34	59	77	83	24					
Space Utilization	104	71	57	73	93	98	40	10				
Devel. & Endow.	54	18	5	18	36	35	0	17	37			
Instruct. Tech.	83	51	23	40	58	69	5	0	14	7		
Work. Devel.	41	15	2	11	25	33	1	20	39	1	15	

Notes.

[a] McNemar-Bowker test is computed with six degrees of freedom.

$\chi^2 \geq 18.7$	Significant at $p \leq .001$, with Bonferonni adjustment
$18.7 > \chi^2 \geq 14.4$	Significant at $p \leq .01$, with Bonferonni adjustment
$14.4 > \chi^2 \geq 11.3$	Significant at $p \leq .05$, with Bonferonni adjustment
$11.3 > \chi^2$	Not statistically significant with Bonferroni Adjustment

Findings: The Keyholder Supplement

In addition to collecting information from keyholders mirroring that gathered from campus chief executives, additional goals established for the KHS, in order of precedence, included: (a) quantifying the extent of their participation in data consortia, exchanges, or benchmarking programs, (b) identifying the types of data shared with outside entities, and (c) understanding the perceived benefit of participation in benchmarking and consortial activities.

Were true keyholders successfully identified? NCES does not disclose keyholder information. Therefore, identifying appropriate respondents was a challenge. In the invited sample, 76% of respondents reported that they were their institution's true keyholder. Because it was possible in convenience sample recruitment materials to explicitly request that only true keyholders participate, the proportion of true keyholders in the convenience sample was higher, at 88%. Chi-square tests of independence indicated that the observed difference was statistically significant, $\chi^2(1) = 11.025$, $p \leq .001$.

Institutional characteristics of true keyholders. The majority (63%) of true keyholders represented 4-year institutions. A smaller number (32%) represented the 2-year sector. Only 5% of true keyholders represented the less than 2-year sector, an artifact of the sampling strategy employed in this study. Respondents were evenly divided between the public and private, not-for-profit sectors (~44%), with 11% of respondents representing the private, for-profit sector.

As was the case in the CES, respondents to the KHS survey came from institutions that tended to be somewhat larger than the sampling frame from which they were drawn: median headcount was 2990 (compared to 1958 in the degree-granting universe), with the 25th and 75th percentiles represented by 1106 (versus 686) students and 8319 (versus 6183) students, respectively.

Roles of true keyholders. Most true keyholders reported that they worked in their campus IR or Assessment office (72%). Other frequently reported roles were that of Registrar (10%) or Chief Executive Officer (8%).

Receipt and Review of the DFR by true keyholders. Eighty-six percent of keyholders reported having received their DFR. That this number did not approach 100% may have been a function of the way the question was put to respondents, not their true experience: the question asked whether respondents had received the *DFR*, when, technically, keyholders received an *e-mail link* to the DFR. Among those who reported receiving their DFR, 94% reported reviewing it. The most frequently reported reason for not reviewing the DFR was "lack of time" ($n = 12$).

Diffusion of the DFR on campus by true keyholders. Respondents who reported having received the DFR were asked about whether they shared it with selected parties. Table 5, below, details true keyholders' experiences with the diffusion of the Data Feedback Report.

Like the parallel analysis conducted on CES data, the results of the KHS survey suggest that the DFR enjoys relatively limited diffusion on campus beyond chief/senior executives and institutional research staff. Much of that diffusion is also passive: only 19% of keyholders report engaging in a conversation with the CEO about the DFR by either leading or joining a review of its contents. The rate of active diffusion by keyholders peaks at 23%, when the DFR shared with other (non-CEO) senior executives.

Keyholders reported they were involved in virtually no diffusion of the DFR off-campus.

Table 5. True Keyholders' Experiences with Data Feedback Report Diffusion

	Respondent reported having …			
Venue, Sorted in Order of Presentation	No Use	Forwarded	Joined a review	Led a review
Accreditors	92%	4%	2%	1%
State/Local Officials	98%	< 1%	< 1%	< 1%
Regents/Visitors	89%	6%	2%	3%
President/Chief Executive	26%	56%	7%	12%
Other Senior Executives	26%	52%	10%	13%
Direct Reports	51%	31%	5%	14%
Campus Governance	79%	13%	3%	5%
IR Staff	51%	31%	5%	14%
Campus Media	86%	10%	2%	3%

Notes.
Totals may not sum to 100% due to rounding error.

Keyholders and institutional research staff interviewed as part of the KHS generally spoke about the DFR in positive terms. Among respondents representing smaller or less well-resource institutions, a frequent subtext was that although they did not make as much use of the DFR as they might like, there was general agreement that the report and its related tools (the Executive Peer Tool, Peer Analysis System, and Data Center) were important. Among respondents representing larger institutions, the DFR and its related tolls were acknowledged as a resource, but one that had been replaced by other services (e.g., the AAUDE Data Warehouse).

A dissenting opinion emerged in an interview I conducted with the Director of Institutional Research at an urban public institution in the Northeast. Expressing serious concerns about the DFR, the respondent reported that it was "not policy-related" and, because of a lack of context, the report was "potentially inflammatory." The respondent went on to suggest that the DFR

could subvert existing strategic planning efforts, making it more likely that decision-makers would "get distracted by something that is not relevant, or about which you can do nothing." Not surprisingly, the respondent reporting not using the DFR, and indicated that if possible, would suppress it in its entirety. Doing so would prevent this respondent's greatest fear: "that some legislator would get a hold of it."

When pressed for more detail about the DFR lacking appropriate context, the respondent suggested that it was due in part to the group of schools to which the institution was being compared, despite having used the DFR's custom comparison group feature in the last three collection cycles. From the perspective of this respondent, the appropriate peer group could vary widely, depending upon the metric in question.

This *particular* view was not widely mentioned among respondents (although it is found in the peering literature, see Appendix A). However, it did relate to a larger theme voiced by others interviewed as part of the KHS: in a peer comparison tool, the group to which an institution is compared is paramount. For example, a respondent representing a health-focused, for-profit institution reported customizing his institution's peer group because campus executives were interested only in schools focused on the same discipline and accredited by the same body. From the perspective of the respondent, a narrowly tailored peer group resulted in a DFR that was both relevant to decision-makers and compelling to a broader audience, including administrators, faculty, and staff. Similarly, a Director of Institutional Research at a community college in California reported that, although the DFR was interesting in that it could provide national context, decisions at a campus and system level were made on the basis of comparisons between his institution and other California community colleges.

Because the group of institutions to which an institution is compared is central to the utility of any peer comparison tool, this topic was the subject of further investigation. That investigation proceeded from the assumption that, for the vast majority of institutions that do not choose to customize their DFR peer group, the accuracy (or lack thereof) of NCES's automatic peer grouping process might influence end-users' perceptions of the DFR's usefulness. In the worst case scenario, a naïve campus executive (or keyholder) might receive their institution's DFR, see a peer group they believe to be irrelevant to their needs, and conclude it is not worth their continued attention. More details about this investigation are presented in Appendix A.

Institutions' participation in consortia and benchmarking activities. After describing their use of the DFR, respondents were asked about their participation in two activities which had been conceived of as being quite distinct:

- Membership in data consortia, defined as organizations where: "members share existing institutional data collections with the intention of sharing results," and

- Benchmarking, defined as working with one or more organizations where: "you share your data with an association, organization, or company and, in return, receive one or more peer comparisons."

Based upon lists of consortia and benchmarking services provided by respondents, it became evident that the definitions above did not provide them enough information to consistently discriminate between the two activities. As such, responses about consortia and benchmarking participation have been combined where possible.

Across all sectors, the rate of participation in some form of data sharing activity was 55%. As can be seen in Table 6, below, some sectors may be disproportionately more likely to participate in these activities than their peers (e.g., 4-year publics), while others are disproportionately less likely (e.g., 4-year private for-profits and non-public 2-years), $\chi^2(5) = 30.040$, $p \leq .001$. Collapsing institutions across level or control suggests that this apparent disproportionality may have been largely driven by control: chi-square tests of independence across levels were not statistically significant, $\chi^2(1) = 2.812$, $p > .05$, while those for control were, $\chi^2(2) = 25.020$, $p \leq .001$.

Table 6. Rate of Participation in Data Sharing Activities, by sector

Sector	Rate
4-year, public	66%
4-year, private not-for-profit	60%
4-year, private for-profit	30%
2-year, public	59%
2-year, private not-for-profit	35%
2-year, private for-profit	14%

Notes.
Because results above are not sector representative, caution should be used in drawing inferences from the above table. Responses from the less than two year sector are supressed due to low response rates.

Barriers to participation in consortia and benchmarking activities. Respondents who had elected to not participate in either benchmarking arrangements or data consortia were asked about the factors which influenced their non-participation. Response options included: (a) a lack of information about available data sharing activities, (b) a belief they already had access to the information available through data sharing, (c) a lack of staff to manage data sharing activities, (d) concerns about data confidentiality, (e) a lack of comparators or consortia, and (f) a belief data sharing would be too costly.

Table 7, below, summarizes respondents' reasons for not participating in data sharing activities. Lack of staff to support data sharing activities, perceived costs, and lack of appropriate venues for comparison or sharing are among the most frequent impediments to accessing peer comparison data.

Table 7. Barriers to Participation in Data Sharing Arrangements

Barrier	% Reporting
Lack of information	43%
Already have access to data	24%
Lack of staff	67%
Concerns about confidentiality	40%
Lack of comparators or consortia	51%
Too costly	53%

In what specific data sharing arrangements are institutions participating? Respondents were asked to list the specific data consortia to which they belonged and benchmarking projects with which they were engaged. Table 8, below, lists those specific arrangements that were most frequently mentioned. Not included there are reports of participation involving membership in state systems of higher education (n = 39), state independent college associations (n = 25), and various regional groups (n = 13).

Table 8. Specific Data Sharing Arrangements Identified by Respondents, by Frequency

Arrangement	Frequency
Consortium on Student Retention Data Exchange	43
National Community College Benchmarking Project	26
Higher Education Data Sharing Consortium	26
National Survey of Student Engagement (Indiana University) Family [a]	25
Association of American Universities Data Exchange	17
University of Delaware Study of Costs	15
Council of Independent Colleges Key and Financial Indicators Toolkits	14
Community College Survey of Student Engagement (University of Texas)	13
Council for Christian Colleges and Universities	13
American Association of University Professors Salary Survey	10
College and University Professional Association for Human Resources Compensation Survey	10

Notes.
[a] Includes Beginning College Survey of Student Engagement, College Student Experiences Questionnaire, Faculty Survey of Student Engagement, and Law School Survey of Student Engagement.

Of those arrangements listed in Table 8, above, two were selected for additional description: the Consortium on Student Retention Data Exchange (CSRDE), by virtue of the frequency with which it was mentioned, and the National Community College Benchmarking Project (NCCBP), because of the unique niche it seeks to fill. For more information about these two projects, see Appendix B.

Frequency and purpose of data sharing. After identifying the consortial/benchmarking projects with which they were involved, respondents were asked to indicate why they chose to share information on a variety of metrics. Reasons included: (a) to guide resource allocation, (b) to assist strategic planning, (c) for the purpose of peer comparison, (d) to share with on-campus stakeholders, (e) to share with off-campus stakeholders, and (f) for accreditation purposes.

Table 9, below, summarizes their responses and, by virtue of the "not applicable, data not shared" option, provides information about the frequency with which various types of data are shared due to consortial membership or benchmarking. Of all the metrics listed, student persistence appears to be paramount: it is most frequently shared, most frequently used in strategic planning and peer comparison, and most frequently shared both on- and off-campus and with accreditors.

Table 9. Purpose of Data Sharing, by Metric

		Percent of Respondents Indicating Purpose[a]					
Metric, Sorted by Rate of Sharing	N/A, Data Not Shared	Resource Allocation	Strategic Planning	Peer Comparison	Share On-Campus Stakeholders	Share Off-Campus Stakeholders	Accreditation
Student persistence and attrition	8%	31%	54%	67%	53%	35%	39%
Student cost of attendance and aid	13%	33%	48%	60%	47%	33%	31%
Student recruitment and admissions	16%	31%	49%	56%	46%	26%	31%
Core revenues and expenses	17%	39%	48%	54%	39%	25%	30%
Faculty diversity	19%	18%	42%	56%	43%	26%	36%
Faculty salaries and benefits	21%	38%	44%	60%	40%	19%	21%
Faculty workload and productivity	28%	34%	40%	48%	38%	19%	30%
Development and endowment	30%	31%	46%	43%	37%	27%	25%
Graduate job placement and workforce devel.	31%	18%	40%	44%	39%	31%	37%
Physical plant and capital assets	31%	35%	42%	43%	31%	22%	28%
Instructional technology	34%	37%	47%	43%	36%	22%	26%
Space utilization	41%	33%	38%	35%	27%	18%	18%

Notes.

[a] Because multiple responses were allowed, percentages exceed 100%

The perceived benefits of data sharing. Respondents who indicated they participated in consortia or benchmarking activities were asked to respond to a series of questions regarding the benefits of doing so, based upon possible reasons for participation previously identified by Sapp (1996). They included the ability to: (a) easily benchmark against relevant peers, (b) use one's own metrics, (c) get results back quickly, (d) ask more specific questions, (e) have access to "sensitive" peer data that would not otherwise be shared, (f) receive technical support from others, (g) build professional networks, (h) increase efficiency of research operations, and (i) build trust with other institutions.

Table 10. Perceived Benefits of Data Sharing Arrangements

Benefit, Sorted by Order of Presentation	Unimportant	Somewhat Unimportant	Somewhat Important	Important
Benchmarking	1%	2%	16%	81%
Use own metrics	2%	4%	22%	72%
Get results quickly	3%	10%	28%	60%
Ask specific questions	3%	6%	33%	58%
Access "sensitive" peer data	5%	11%	35%	49%
Receive technical support	4%	20%	38%	38%
Build professional networks	4%	18%	34%	43%
Increase efficiency	8%	22%	34%	36%
Build trust	7%	17%	32%	44%

Notes.

Row totals may not sum to 100% due to rounding error.

Findings: The Data Providers Survey

The goals established for DPS, in order of precedence, included: (a) conducting a census of providers of comparative data, and (b) identifying the types of data made available by providers of comparative data and its method of dissemination, (c) determining the source(s) of data made available by providers.

Eligibility to complete DPS. The first question posed to each DPS respondent was designed to determine whether the organization which she or he represented was a likely provider of peer comparison data. As such, each interview began with the following eligibility screening question:

> Does your organization provide data, tools, or products to colleges or universities specifically designed to allow them to compare themselves to relevant peers?
>
> Examples might include, but are not limited to:
> (a) Conducting a student survey at a group of institutions and reporting institution-by-institution results to participating schools,
> (b) Gathering institution-level data from a group of colleges and synthesizing it into a comparative report,
> (c) Benchmarking an institution against a selected group on the basis of data already collected, or
> (d) Providing a Web-based tool allowing an institution to develop its own comparisons.

More than 63% (n = 131) of the 206 respondents to the DPS indicated that their organization did, in fact, provide some form of comparative data. Ineligible respondents were thanked for their participation and exited from the survey.

Characteristics of data providers. To understand the types of organizations that provide peer comparison data, respondents were asked to classify the structure of their organization by selecting from among the following five options: (a) a consortia of institutions or states, (b) a non-profit association, (c) a for-profit association, (d) a governmental agency, or (e) other, with a free-text box to provide additional information. The overwhelming majority (73%) of respondents were non-profit organizations of some sort, including associations and individual universities. Approximately 23% of respondents represented consortia of institutions or states. The remaining five cases included for-profit corporations (n = 3), accreditors (n = 2), and governmental agencies (n = 1). Organizations were also generally well-established: 88% had been operating for more than 20 years.

Although all respondents represented organizations that, to some extent, provided comparison data, rarely was that their primary purpose: only 11% of respondents indicated that peer comparison was their primary activity. Irrespective of whether the provision of peer comparison services was a key business practice of responding organizations, data suggests that their efforts are relatively longstanding: 47% of providers reported having offered peer comparison products for more than 20 years, and only 16% had entered the marketplace in the last five years.

Fees. Respondents were asked to identify whether fees were associated with membership in their organization, for participation in peer comparison projects, and, if applicable, for *ad hoc* or "special" analyses. Almost all providers (86%) charged some form of membership fee, with the few exceptions typically being research programs affiliated with a university. By comparison, a relatively small number (23%) charged additional fees for peer comparison products. Of those providers who completed *ad hoc* analyses for the members above and beyond the standard peer comparison products they already provided, 40% did so for an additional charge.

Types of comparisons offered. After providing information about their agency and its history, respondents were asked whether they made comparisons available on the list of metrics previously presented to campus chief executives and keyholders. Table 11, below, summarizes their responses. Perhaps not surprisingly, the frequencies with which data providers reported providing comparative products focused on each of our metrics closely mirrored chief executives' ratings of those metrics' importance. What cannot be determined, of course, is whether the availability of data is a *response to* demand from campus decision-makers, or whether, in actuality, the availability of data is a *creator of* demand.

Table 11. Types of Comparison Data Offered by Providers

Metric, Sorted by Frequency of Offer	Percent of Providers Offering
Student persistence and attrition	59%
Student cost of attendance and aid	56%
Student recruitment and admissions	50%
Core revenues and expenses	48%
Faculty salaries and benefits	41%
Faculty diversity	33%
Development and endowment	32%
Physical plant and capital assets	29%
Instructional technology	25%
Graduate job placement	24%
Faculty workload and productivity	22%
Space utilization	20%

Source of comparison data. Next, respondents were asked to select the source(s) of data they used in their peer comparison products from among a set of five options: (a) their own original data collection, (b) the data collections of other organization (e.g., rather than conducting their own survey of endowments, for example, an alternative data provider might instruct members to participate in the NACUBO Endowment Study, copying them on all collection materials), (c) the data collection of state higher education authorities, (d) IPEDS survey components, and (e) other federal data collections (e.g., the National Science Foundation's Survey of Science and Engineering Research Facilities). Results are summarized below, in Table 12.

As can be seen in Table 12, most providers collect their own data. However, the repackaging of IPEDS data runs a close second on all metrics for which IPEDS collects data.

Table 12. Sources of Data for Peer Comparison Products

	Data Source, Among Those Providing Comparison[a]				
Metric, Sorted by Use of IPEDS Data	Original Collection	Other's Collection	State Collection	IPEDS Data	Other Federal Data
Core revenues and expenses	59%	12%	4%	39%	2%
Student persistence and attrition	68%	17%	17%	34%	11%
Faculty salaries	72%	15%	2%	33%	2%
Student cost of attendance and aid	68%	15%	16%	32%	5%
Development and endowment	65%	21%	3%	26%	0%
Faculty diversity	81%	8%	6%	26%	0%
Student recruitment and admissions	83%	17%	2%	22%	4%
Physical plant and capital assets	77%	17%	7%	20%	0%
Graduate job placement[b]	68%	24%	4%	12%	8%
Space utilization[b]	85%	20%	0%	10%	0%
Faculty workload and productivity[b]	78%	9%	0%	9%	4%
Instructional technology[b]	77%	19%	0%	4%	0%

Notes.

[a] Because multiple responses were allowed, percentages exceed 100%

[b] Data on this metric is not available through an IPEDS survey component.

Granularity of data. After identifying the metrics on which they provided peer comparison products, respondents provided additional information about the level of granularity at which comparisons were presented. Results are summarized below, in Table 13.

Table 13. Granularity of Data Provided by Peer Comparison Products

	Level, Among Those Providing Comparisons		
Metric	Program	Institution	Multiple Institutions
Student persistence and attrition	40%	69%	43%
Student recruitment and admissions	30%	65%	35%
Student cost of attendance and aid	19%	74%	37%
Faculty salaries and benefits	28%	50%	39%
Core revenues and expenses	25%	67%	37%
Faculty workload and productivity	65%	30%	17%
Faculty diversity	50%	53%	19%
Graduate job placement and workforce devel.	40%	48%	32%
Development and endowment	21%	71%	26%
Instructional technology	35%	58%	42%
Physical plant and capital assets	27%	63%	40%
Space utilization	30%	55%	40%

Notes.
Because multiple responses were allowed, row percentages may exceed 100%

Student or faculty unit records. As a follow-up to the question on the DPS related to the granularity of comparison provided to users, respondents were asked about the granularity of data collected on students and faculty. Student unit records were gathered by 16% (n = 21) providers, and generally consisted of demographic information, pre-college characteristics, enrollment data, and course of study. One system (maintained by a disciplinary accreditor) tracked subsequent employment. Faculty unit records were gathered by 12% (n = 16) of respondents, and typically included demographic information, rank, and discipline. In rare instances, salary histories, benefits packages, and non-specific "productivity" data (perhaps publications or extramural funding) were collected.

Method of diffusion. Respondents were asked to identify the ways in which their organization diffused the results of their peer comparison work. Their responses, summarized in Table 14, below, suggest that the bulk of comparison data available to consumers is static. Web tools are less common, and access to raw data even more so.

Table 14. Method of Diffusion Employed by Providers of Peer Comparison Data

	Among Those Providing Comparisons		
Metric	Report (Print or PDF)	Raw Data	Web Tool
Student persistence and attrition	69%	20%	31%
Student recruitment and admissions	69%	22%	34%
Student cost of attendance and aid	69%	21%	29%
Faculty salaries and benefits	61%	24%	35%
Core revenues and expenses	63%	16%	27%
Faculty workload and productivity	52%	17%	43%
Faculty diversity	67%	14%	42%
Graduate job placement and workforce devel.	56%	16%	28%
Development and endowment	76%	18%	24%
Instructional technology	62%	27%	42%
Physical plant and capital assets	70%	23%	30%
Space utilization	50%	20%	35%

Notes.

Because multiple responses were allowed, row percentages may exceed 100%

Access to institutionally-identifying comparisons. Table 15, below, suggests that a user's access to institutionally-identifying data in a peer comparison product is generally (although not exclusively) linked to the extent of his or her participation in a data collection. That said, only slightly more than half (55%) of respondents indicated that they made any form of institutionally-identifying information public to participants. Among non-participants and non-members, institutionally-identifying information is even rarer (at 28% and 13% of respondents, respectively).

Table 15. Level of Access to Peer Comparisons, by Participation Status

Level of Access	Participant	Non-Participant	Non-Member
No Access	2%	30%	56%
Non-Institutionally-Identifying Data	43%	42%	30%
Institutionally-Identifying Data	55%	28%	13%

Notes.
Column totals may not sum to 100% due to rounding error.

Emerging metrics. Finally, respondents to the DPS were asked to list metrics they felt were, or soon would be, important new directions for peer comparisons among their members. Metrics receiving more than one mention are listed below, in Table 16. Metrics related to institutional finance (particularly indicators of "financial viability"), sustainability/environmentalism, and student success were most frequently mentioned by respondents.

Table 16. Emerging Metrics Noted by Two or More Respondents

Metric	Frequency
Finance-related (including investment and endowment)	11
Environmental sustainability	10
Persistence and graduation	9
Student learning	6
Student diversity	4
Workforce development/graduate job placement	4
International issues	4
Admissions and enrollment	2
Student financial aid	2
Remedial course-taking	2

Conclusion

At its inception in 2004, the IPEDS DFR (and related web-based tools) was the only peer comparison tool available to the majority of institutions in the postsecondary universe. Six years later, that is still the case. The DFR is free to all users, covers a wide range of metrics, facilitates comparisons against peer groups of virtually any combination imaginable, and is accessible in a variety of forms. As such, the DFR is positioned to be *the* preeminent source of comparative data nation-wide. The results of this study suggest that the DFR has not yet attained that status, however, and offers some insight as to why this might be. In conclusion:

- While DFR utilization rates are hard to determine, a potential proxy—the rate at which institutions choose to upload custom comparison groups to enhance the utility of their DFR—varies widely by sector. While this rate hovers near 50% for public, 4-year institutions, it falls to approximately 5% among for-profit institutions, overall. Even if utilization rates are twice as large as rate of customization, this suggests a large swath of postsecondary institutions is not using the DFR.

- Campus executives, the target audience for DFRs, report relatively little diffusion of the product beyond the cabinet level. Rates of diffusion off-campus are particularly low.

- Based upon feedback provided by campus executives, the metrics contained on the DFR are appropriate for its purpose. All were rated as "somewhat important" by at least 70% of campus executives, although a clear preference for student-based metrics was detected.

- Although campus executives are most prone to turn to on-campus research staff to get information they need for their day-to-day work, their use of the DFR is on-par with other off-campus data sources.

- Alternative data providers in general—be they benchmarkers or consortia—are enjoying relatively frequent use by institutional research staff, although their use is greatest among public institutions and 4-year, not-for-profits. Institutions report using alternative data providers because they facilitate benchmarking, they allow the use of preferred metrics and the answering of specific questions of interest to the institution, provide access to confidential data, and have a relatively rapid turn-around time (often within the quarter).

- The Consortia for Student Retention Data Exchange (CSRDE) and the National Community College Benchmarking Project (NCCBP) were amongst the most frequently mentioned alternative data providers in this study. More details about their work can be found in Appendix B. CSRDE's prominence parallel's chief executive's most important metric: student persistence and attrition.

- Most alternative data providers are collecting original data from institutions. However, depending upon the metric, IPEDS data is relatively frequently "repackaged" by alternative data providers. Providers that do not explicitly use IPEDS data but instead run their own collections indirectly use IPEDS data, often directing their users to the appropriate IPEDS data element to identify the specific data value to report.

- Although web-based delivery of peer comparison data is on the rise, paper (or PDF) remains the predominant diffusion format.

- Indicators of institutional financial health and institutional sustainability efforts were among the emerging metrics most frequently identified by alternative data providers as being increasingly relevant for their constituents, although indicators of student success remain important.

Appendix A: Peer Grouping

Much has been written about the ways in which institutions come to develop peer groups for the purpose of inter-institutional comparison: Indeed, Hurley's (2002) review cited nearly a dozen pieces authored between 1980 and 2000 that address the topic. In that review, Hurley noted that authors have primarily distinguished peer grouping methodologies along two lines: (a) type, which refers to the nature of the relationship between the focal institution and its comparators, and (b) method, which refers to the way in which peer institutions are selected.

The peer grouping literature generally recognizes four distinct group types: (a) competitor, (b) peer, (c) aspirational, and, depending upon the author, (d) predetermined/jurisdictional (Hurley, 2002). Each type's name implies the nature of the relationship between the focal institution and its group of comparators, and can include institutions that are seen as direct competitors, institutions that are similar on one or more dimensions but may not share a common market, institutions to which the focal institution aspires to become more like, and institutions that are *de facto* or *de jure* peers, such as fellow members of a state system.

Similarly, four methods of peer group determination are commonly cited in the peer grouping literature (Brinkman & Teeter, 1987; Hurley, 2002). Perhaps most common is the *panel* method, in which stakeholders/informants generate a list of an institution's peers based upon their beliefs/knowledge about the focal institution and potential comparators. Both the *threshold* and the *hybrid* method hinge upon identifying, and then weighting based upon their relative importance, variables believed to be important in the peering process. While the threshold method seeks to identify similar peers by determining their "nearness" to pre-defined ranges/categories (e.g., student headcount between 1000-2500, within a given region, confers between 50-100 education degrees yearly), the hybrid method standardizes raw scores on the identified variables and generates a weighted similarity score. Then, for both methods, institutions are ranked by score and the ranked list used to identify peers. The use of *cluster analytic* methods is the most statistically complex and ostensibly objective (Hurley). First, potentially important peering variables are identified by the analyst. Then, statistical software is used to identify, in multivariate space, those colleges or universities most similar to the focal institution.

Currently, NCES employs a version of the *threshold* approach for the generation of automatic peer groups. For 2008-2009, the DFR universe is slated to broken in to 227 groups on the basis of a variety of characteristics, including: (a) control, (b) level, (c) reporting method, (d) degree-granting status, (e) largest CIP, (f) unduplicated headcount, (f) urbanicity, (g) geographic region, and (h) Carnegie classification. The resultant groups include examples such as "Private, not-for-profit, degree-granting institutions in Puerto Rico" and "Theological schools that offer undergraduate degrees and no graduate degrees beyond the Master's." A similar approach was

used for 2007-2008, resulting in 223 automatically generated, mutually-exclusive comparison groups, comprising 6692 institutions.

As noted earlier in this report, the use of customized peer groups remains relatively low (approximately 20% universe-wide). Additionally, there is a relationship between institutional sector and peer group customization: 4-year institutions outpace their 2-year and less-than-2-year counterparts in the rate of customization, as do public institutions when compared to privates (and particularly privates that are for-profit). The concern becomes the relationship between use, utility (presumably promoted by customization), and institutional capacity: Are institutions that are ill-prepared to identify a solid peer group not using the DFR because the group of institutions to which they are being compared is inappropriate? Furthermore, might better automatic peer grouping results in more DFR usage?

The answers to these questions are not discernable from the present study: survey response and rate of DFR use appear to be conflated. However, if we accept on face the argument that better peer grouping can only serve to increase DFR use, then two additional questions arise. First, what can be learned from those institutions that currently customize their peer groups? Second, how might we judge the "goodness" of a given grouping scheme?

To explore these questions, data was extracted from the IPEDS collection system, including the UNITID of every institution that had uploaded a custom comparison group for 2007-2008 and the UNITIDs of every institution it had identified as a peer. Then, data corresponding to each DFR indicator and several institutional characteristics were matched to each UNITID in the file. A similar file was created for the peer groups automatically generated by NCES.

Note: It was assumed at the outset of the analyses presented below that the threshold approach currently employed by NCES is, in some form, the only viable method of peer group determination. Compared to panel methods, the threshold approach is free from obvious subjectivity. (Of course, this method would not be feasible if for no other reason than it would also be impossible for NCES staff to be sufficiently knowledgeable about each institution in the DFR universe to identify potential peers.) The threshold approach is also preferable to the hybrid or cluster analysis methods—for most audiences, at least—because both its methods and results are readily comprehendible and require no statistical training.

<u>Do custom comparison group users employ threshold techniques?</u> Technically, the use of threshold approaches relies upon not only the identification of potentially relevant peering variables but also a scheme for weighting those variables that reflects their perceived importance. As such, it is not possible to replicate any individual institution's threshold process. However, it is possible to determine whether the variables used by NCES to segment the DFR universe into mutually-exclusive group are rigorously employed by custom comparison group users.

To do so, I began by identifying a series of potentially important "thresholding" variables, including: (a) state, (b) OBE region, (c) sector, (d) level, (e) control, (f) locale/urbanicity, (g)

2000 Carnegie class, (h-m) 2005 Carnegie basic and subtypes, (n) core-based statistical area, and (o) 12-month FTE quintiles. Then, for each customized comparison group, a "percent match" statistic was computed, representing the percent of selected peers in that group that shared the same value as the focus institution (that is, they both were in the same state, region, sector, control, and so forth). A total of 1328 groups were evaluated. Presumably, higher values represent greater support for the use of a given variable as a threshold. Heuristically, the following analytic process was employed:

```
for each threshold variable {
          for each group {
                    numpeers = n(members) – 1
                              for each peer {
                                        match = 1 if peer(variable)==focus(variable)
                                        }
                    sum_matches = sum(match)
                    match_statistic = sum_matches / sumpeers
                    }
}
```

Table 17, below, summarizes the mean percent within-group matches on key thresholding variables in custom comparison groups. Based upon the data presented there, sector, level, and control appear to be the most salient thresholding variables employed by custom peer group users, at least among the variables tested. Given the centrality of those variables to the NCES thresholding scheme, high mean percent within-group matches are a "good thing." The same is true, albeit to a lesser extent, of variables representing Carnegie classification.

Variables that can be used to infer spatial relatedness among institutions, including core-based statistical area (CBSA), state, and region, demonstrate a pattern that may suggest interesting consequences for peer grouping for certain sectors. Four-year public institutions, for example, demonstrated relatively little similarity with their peers in terms of CBSA (8%), only slightly more for state (12%), and somewhat more on region (34%). In contrast, four-year for-profit institutions demonstrated substantial similarity in terms of CBSA (46%), state (66%), and region (76%). One tenable explanation of this stark difference by control is that 4-year for-profit institutions are creating "markets" that are relatively local, while 4-year public institutions presume to draw from a larger geographic region. Two-year institutions' relatively high similarity on state (47%) may signify the use of jurisdictional peers suggested by Brinkman and Teeter (1987), and is consistent with comments offered by presidents of public, 2-year institutions that their only relevant peer group is the one made up of other community colleges in their state.

The use of spatial variables in thresholding is, unfortunately, not without its problems. Chief among them is an uneven sector-wise distribution of institutions by CBSA, state, and region resulting in ill-sized groups. In 2007, the average size of an NCES-generated peer group was approximately 30, with no group consisting of fewer than 5 institutions. In a sector × region matrix (excluding administrative units and service academies), 28 of the 81 cells have fewer than 30 institutions and five cells have fewer than five institutions. In a sector × state matrix

(excluding administrative units and outlying areas), 318 of the 459 cells have fewer than 30 institutions and 122 cells have fewer than five institutions. Not surprisingly, the problem of ill-sized groups becomes worse when using CBSA. Initially, 324 institutions are found in areas too rural to be included in a CBSA. Of those that remain, and aggregating across all sectors, 766 of 810 CBSA's represented in 2007 had fewer than 30 institutions, and 558 had fewer than five institutions. Adding other variables to a thresholding scheme involving spatial variables further complicates matters: Indeed, only a small subset of NCES automatically generated peer groups include any spatial variable and, in 2008, regions had to be collapsed to create well-sized groups.

Despite the complications involved, however, it seems that some institutions—most notably for-profits—*are* thresholding using spatial variables, sometimes at very fine levels of detail.

<u>What other thresholds might be beneficial?</u> Above, the notion of using spatial variables for thresholding, at least in certain sectors, is considered. Additional threshold variables may be useful for unique institutional types. As a group, cosmetology schools, which numbered more than 1000 in 2007-2008, is one such example. In the current NCES scheme, cosmetology schools are subdivided strictly by size (specifically, they are placed in size order and then broken into groups of approximately 30). In conversations with knowledgeable practitioners, the notion of first subdividing cosmetology schools by separating "corporate" from non-corporate schools was suggested as mechanism to control for varying levels of support and autonomy between institutions.

While IPEDS does include a variable [f1systyp] that asks whether an "institution is part of a System, Governing Board, or Corporate Structure," an inspection of the use of that variable has suggested that it is inconsistently used. According to IPEDS Helpdesk Staff, RTI's efforts to ensure greater consistency on f1systyp have been substantial but yielded little success. If it is the case, as practitioners have suggested, that distinguishing corporate from non-corporate cosmetology schools is a meaningful distinction, then a reworded question on the appropriate IPEDS survey component (IC) would seem to be indicated. It may be that a data element currently maintained in FSA's PEPS that helps organize details related to institutional ownership would provide a possible template for f1systyp's ultimate revision.

<u>How might comparison group efficiency be evaluated?</u> The identification of one or more metrics for assessing the efficiency of alternative grouping schemas is important in identifying which (if any) schemas are quantitatively "better" than others. To be sure, there is a qualitative method for doing so: analyst review. Analyst review is problematic, though, because of its subjectivity and its demand that the analyst be incredibly knowledgeable about the DFR universe. As such, quantitative techniques are likely to be more desirable.

I suggest that the appropriate operationalization of a grouping schema's efficiency is its capacity to minimize within-group variability and maximize between-group variability. Within the analysis of variance framework, this ratio is analogous to the intraclass correlation coefficient

(ICC) in a one-way random effects model. Unfortunately, this results in a measure that can be applied to only a single variable or, in this case, a single DFR indicator.

The notion that peer group efficiency is metric-dependent is not without precedent. Anecdotally, an IR director surveyed as part of this work insisted that her institution had a variety of peers, depending upon the "thing" being contrasted. Published work by Whiteley and Stage (1987) further supports the notion of issue-based peering. Indeed, what appears to be a bother may, in fact, be of interest (if only for theoretical purposes) to compare or contrast the interactions between sets of variables, within-cluster variability, and different peering mechanisms.

Generally, though, some omnibus technique for characterizing the efficiency would seem to be helpful. As a result, the weighted average ICC, presented below, is suggested in the case when two schemes are being tested on identical populations. For comparisons on *k* indicator variables:

$$\rho_{wa} = \sum_{i=1}^{k} \frac{ICC_k \times \sigma_k^2}{k}$$

In an effort to reward schemes that produce high ICCs on indicators with high variability, this approach "weights" each indicator's ICC by its variance. The resulting scores are then averaged across the set of indicators being evaluated. Certainly, alternative conceptualizations or potential refinements could be considered. Assigning "importance weights" to each indicator could, for example, allow the analyst to develop an index that rewarded schemes that effectively partitioning variance on a subset of key indicators, while minimizing the impact of less important ones.

On the next page, in Table 18, the ICCs associated with the NCES automatically-generated peers and institution's own custom comparison groups are detailed for nine DFR metrics. Statistically significant differences between ICCs (determined by non-overlapping 95% confidence intervals) are starred. Findings were mixed, although it is possible to speculate about observed differences. The lack of statistically significant difference in ICCs related to size and faculty salaries is not wholly surprising, to the extent that both institutions and NCES use size and Carnegie classification (and sometimes region) to determine peer groups. Institutions may be more likely to select peers with similar student characteristics, including race/ethnicity, need, and propensity to graduate, yielding higher ICCs for institutional comparison groups on those dimensions. As a reminder, the weighted average ICC *cannot* be computed here, because not all institutions are represented in the institutionally-generated peer groups.

Table 17. Mean Percent Within-Group Match on Key Thresholding Variables in Custom Comparison Groups

	4-Year			2-Year			Less-Than 2-Year		
	Public	N-F-P	F-P	Public	N-F-P	F-P	Public	N-F-P	F-P
Variable description [variable name]	n = 349	n = 578	n = 63	n = 282	n = 18	n = 9	n = 2	n = 1	n = 22
State [fips]	12%	26%	66%	47%	28%	57%	74%	4%	39%
Region [obereg]	34%	51%	76%	66%	46%	66%	74%	29%	56%
Sector [sector]	86%	90%	49%	89%	50%	72%	45%	0%	79%
Instructional level [level]	91%	94%	84%	90%	63%	76%	45%	0%	83%
Control [control]	90%	90%	55%	94%	76%	91%	92%	96%	89%
Urbanicity [locale]	19%	21%	35%	20%	20%	23%	28%	8%	24%
2000 Carnegie classification[a] [carnegie]	64%	54%	22%	90%	70%	66%	n/a	n/a	n/a
2005 Carnegie basic classification[b] [ccbasic]	64%	53%	24%	88%	64%	67%	n/a	n/a	n/a
Core-based Statistical Area [cbsa]	8%	14%	46%	12%	22%	26%	28%	8%	19%
Quintiles of 12 Month FTE [fte12mn]	46%	36%	19%	38%	24%	21%	29%	13%	25%

Notes.

[a] Excludes institutions (n = 106) not classified by Carnegie in 2000

[b] Excludes institutions (n = 50) not classified by Carnegie in 2005

[c] Excludes institutions (n = 74) not located in a CBSA

Table 18. A Comparison of the Clustering Efficiency of NCES Automated and Institutional Custom (IC) Peer Groups

Metric	Variable Information	NCES ICC (SE) (n=223)	IC ICC (SE) (n=1328)	Sig.
12-Month FTE	fte12mn	.624 (.024)	.609 (.012)	
White, non-Hispanic enrollment	pctenrwh	.283 (.022)	.351 (.012)	*
Tuition and fees	tufeyr, if rptmth = 1	.801 (.021)	.840 (.007)	
Average federal loans, FT-FT UG's	fgrnt_a	.196 (.018)	.361 (.012)	*
Overall graduation rate, FT-FT UG's	grrttot	.483 (.026)	.704 (.010)	*
Bachelors degrees awarded	basdeg, if iclevel = 1	.794 (.025)	.677 (.012)	*
Instructional expenses/FTE	f1instft, if control = 1	.126 (.024)	.491 (.018)	*
	f2instft, if control = 2	.000 (.007)	.018 (.004)	
	f3instft, if control = 3	.226 (.032)	.000 (.009)	*
FTE staff, instruction/research/service	irpsstf	.785 (.017)	.667 (.011)	*
Average faculty salary, all ranks	saltotl	.627 (.029)	.654 (.011)	

Appendix B: Highlighted Data Sharing Arrangements

As noted earlier in this report, two data sharing arrangements were selected for additional description in this report: the Consortium for Student Retention Data Exchange (CSRDE), due to the frequency with which respondents reported participating, and the National Community College Benchmarking Project (NCCBP), due to the unique niche which the project seeks to fill.

CSRDE[9]. Operating since 1994 and today comprised of more than 660 members, CSRDE is one of the largest (if not *the* largest) consortia of institutions operating in the United States and Canada. Since its inception, it has been housed at the University of Oklahoma. As its name implies, its focus is simple: sharing data on retention and graduation from both two-year and four-year postsecondary institutions.

For four year institutions, data is collected on three cohorts of students: (a) first-time, full-time, degree-seeking first-year undergraduates, (b) first-time, full-time, degree-seeking first-year undergraduates in STEM disciplines, and (c) transfers from community colleges. For two-year institutions, data is collected on two student cohorts: (a) first-time, full-time students, and (b) first-time, part-time students.

The collection itself is relatively straightforward. Information is submitted electronically via a spreadsheet to CSRDE on a yearly basis (see a facsimile in Exhibit 1). Data elements include: (a) total headcount, (b) undergraduate headcount by enrollment intensity and degree-seeking status, (c) student characteristics, including average ACT and SAT scores, as well as percent of students over 24; living on campus; who are racial and ethnic minorities; who were in the top 10%, 25%, and 50% of their graduating high school class; who are receiving federal grants, and (d) year-to-year continuation or graduation rates for cohort members, up to 200% of total time to degree. Data on each year's cohort is collected both in the aggregate (the "total" cohort), and is then disaggregated by sex and race/ethnicity (here, Black, Hispanic, Asian, American Indian, White, and not reported).

Data on STEM retention is gathered in the same manner. As is the case with all completions studies that focus on a subset of disciplines, an analyst's criteria for cohort membership—usually established by identifying some number of CIP codes that are believed to encompass the discipline in question—will have consequences for a study's findings. In the case of the CSRDE STEM retention study, the following CIP 2000 codes are included: (a) 03.xxxx (Natural Resources and Conservation), (b) 11.xxxx (Computer and Information Sciences and Support Services, (c) 14.xxxx (Engineering), (d) 15.xxxx (Engineering Technology/Technicians) , (e) 26.xxxx (Biological and Biomedical Sciences), (f) 27.xxxx (Mathematics and Statistics), (g) 40.xxxx (Physical Sciences), (h) 01.0000 (Agriculture General), (i) 01.0801 (Agricultural and

[9] See http://csrde.ou.edu/, retrieved July 10, 2009.

Extension Education Services), (j) 01.09-01.999 (Animal Sciences), (k) 30.1901 (Nutrition Sciences), and (i) 30.2401 (Neuroscience).

Reporting typically lags data submission by three months. Four "universe" reports are available, including: (a) a summary of four-year institutions, (b) a summary of two-year institutions, (c) a summary of STEM retention at four-year institutions, and (d) a summary of the community college transfer report. Each institution is also entitled to one customized peer report per survey completed. All reports are available in print/PDF.

CSRDE does operate a web-based interface to its data warehouse, known as the Quick Query System. An institution's level of access to the system depends upon the level of membership they have purchased with CSRDE. At Level 1, members use the system *only* to upload their customized peer group and immediately extract their complementary peer report(s), as opposed to sending their peer information to CSRDE and waiting for the report to be manually generated. At Level 2, members have access to the Level 1 peer reporting system, and can also use the system to subset the larger CSRDE universe and generate summary reports for groups of institutions by institutional characteristics. No limit is imposed on the number of summary reports that institutions may generate at this level. At Level 3, institutions have access to the Level 2 summary reporting system, and also have unlimited access to the Level 1 peer reporting system.

CSRDE offers other benefits to its members in addition to reporting the results of its data collections. This includes the yearly National Symposium on Student Retention and monthly retention-related webinars. Additional fees are charged for these opportunities, although Level 2 and Level 3 members receive three and six complementary webinars, respectively.

CSRDE's pricing structure is not publicly available. However, inquiries among practitioners suggested that Level 1 membership is $475.

NCCBP[10]. The NCCBP is a project of Johnson County (KS) Community College's National Higher Education Benchmarking Institute (which also operates the so-called *Kansas Study*, a study of instructional costs and productivity of two-year institutions). Since its inception in 2004, the NCCBP has grown, in 2008, to cover 188 institutions. Reacting to "increasing pressures to demonstrate effectiveness to federal, state, and accrediting agencies,[11]" the NCCBP seeks to provide participating institutions a mechanism for benchmarking against relevant peers and, in a limited manner, identify potential best practices.

Like CSRDE, NCCBP collects data from its participants electronically, via spreadsheet. Collection materials become available in March and are due in June. The list of data elements collected by NCCBP is quite substantial (see Exhibit 2). From them, twenty-five distinct

[10] See http://www.nccbp.org/, retrieved July 10, 2009.
[11] See http://www.nccbp.org/History, retrieved July 10, 2009.

benchmarks are derived, including persistence and graduation rates, success of students enrolled in developmental course-work, and several service-area specific rates, including market penetration, community participation, and service to racial/ethnic minority populations.

Dissemination of NCCBP results begins in October. According to the NCCBP website, four types of reports are available, including: (a) an "aggregate" report that includes an institution's scores on each metric, along with translations of benchmark scores to percentile ranks, (b) a "subscriber directory," which provides a summary of the institutional characteristics of participating schools, (c) an unlimited number of peer comparisons (although peer data is stripped of institutional identifiers) that provide reported values for the focus institution, the comparison group, and the 10th, 50th, and 90th percentiles of the NCCBP "universe," and (d) a "best practices" report, in which institutions that have placed in the 80th percentile on a given metric and have waived their option of anonymity are listed.

Neither aggregate reports nor peer reports are available to non-participants, nor are they to be made available by participants in publically-available formats. Nonetheless, several examples of both aggregate and peer reports *have* been posted on institutional web-sites and are available for inspection[12]. Exhibit 3 (included) is a standard NCCBP institutional report, while Exhibit 4 (included) is a customized peer report. A review of both report formats reveal that they are substantively as described on the NCCBP website.

Unlike CSRDE, NCCBP has no participation "levels." All institutions remit a $1000 fee for participation, and, after doing so, have access to all study materials in which they participated (that is, aggregate or peer data is available only for those metrics an institution provided).

[12]See http://www.google.com/#hl=en&q=NCCBP+Web+filetype%3Apdf, last accessed July 10, 2009.