USNWR College Rankings Reexamined

Thomas J. Webster, (E-mail: twebster@pace.edu), Pace University
Rosette J. Mare, Stony Brook University

ABSTRACT

This paper extends Webster's [2001] analysis of the accuracy of the weighting scheme utilized by U.S. News & World Report (USNWR) to rank colleges and universities according to "widely accepted indicators of national excellence," which he found to be plagued by severe and pervasive multicollinearity. As in the Webster study, we employ principal component analysis to assess the relative contributions of thirteen criteria used by USNWR in 2004 to rank "top schools" in the national university category. Although USNWR continues to assign the greatest weight to peer assessment, this study confirms Webster's findings that average SAT/ACT scores of enrolled students is the most significant ranking criterion. This paper also extends Webster's study by examining the reliability of the USNWR rankings, which have come under repeated criticism for their lack of consistency. When compared with simulations generated from an estimated principal component regression model, the 2004 USNWR rankings are found to be increasingly more unreliable for lower ranked institutions. The source of this inconsistency appears to be peer assessment, which is the only subjective criterion used in the USNWR ranking methodology. This suggests that the rankings might be improved by lowering (or removing entirely) the relative contribution of peer assessment from the USNWR ranking methodology.

INTRODUCTION

There is an ongoing and heated debate about the efficacy and usefulness of the *U.S. News & World Report* (*USNWR*) tier rankings of colleges and universities. These rankings, which have been published by *USNWR* since 1981, purportedly identify the best and worst colleges and universities in the United States. *USNWR* has repeatedly argued that the college rankings should not be used exclusively to evaluate the relative quality of each institution included in the rankings. Rather, they argue, prospective students should judge each institution on its relative merits. On its website, *USNWR* asserts:

The U.S. News rankings of colleges and universities provide an excellent starting point for families comparing colleges because they offer an opportunity to judge the relative quality of the educational experience at schools according to widely accepted indicators of excellence. But many other factors that can't be measured should figure in your decision, including the school's cost, the availability of aid, course offerings, the feel of campus life, and the setting and geographic location.

So how should you use our ranking tables? Study the data that accompany the actual rankings. The tables are a source of highly useful information about colleges that is otherwise hard to obtain and which will help you narrow your search to a small number of colleges that are a good fit."

If the real value of the rankings is the data contained in the tables, then why continue to rank these institutions? Notwithstanding the importance of the rankings to *USNWR* sales, they have become important to prospective students because it makes the college search process more efficient and less costly.[1] The rankings are also used by some college administrators (when it suits them) in their marketing and fund-raising campaigns. Because the *USNWR* rankings influence prospective students' perceptions of their institution, they are used by administrators to define the institution's market niche, enrollment targets and operating budgets (see, for example, Carter [1988], Crissey [1997], Garigliano [1997], Gilley [1992], Glass [1997], Gleick [1995], Graham and Diamond [1999], Kirk and Corcoran [1995], Marchung [1998], Morse and Gilbert [1995], and Schatz [1993]). The rankings are even used by prospective employers to allocate limited college recruitment budgets. Whether you love them or hate them, the rankings have become an integral part of the college landscape.

Because of their increasing importance to students, administrators, corporate recruiters and others, the *USNWR* college rankings have spawned considerable controversy. At the heart of the debate is the apparently subjective nature of the ranking process. In particular, *USNWR* has come under heavy criticism for its arbitrary selection, and weighting, of ranking criteria.

...educators criticize the entire system used by the magazine. Alan J. Stone, president of Alma College, said the data collection and weighting 'simply missed the mark.' Mr. Stone...said the magazine's methodology...'was so subjective, it is ridiculous.' [Glass, p. 93]

'Getting rid of the U.S. News' rankings would be one of the best things that could happen to higher education,' said Michael Aiken, the University of Illinois chancellor, summing up what nearly every other prominent education expert believes. The rankings, they say, are based on shoddy methodology. To them, ranking colleges is like ranking music, an inherently flawed idea. [Glass, p. 94]

College administrators may not like the rankings, but they are forced to deal with them because of the profound impact that a change in the rankings could have on the institution's operations. "One mid-Atlantic college's admissions officer says that his college president will probably fire him if they fall in the rankings this year." [Glass, p. 94] Understandably, the pressure to do well has resulted in questionable reporting practices by ranked institutions.

...the U.S. News editors know that some colleges cook their numbers, while administrators mistrust any system that seeks to quantify their institution's intangible strengths–they simply need each other. From a business standpoint, the magazine obviously needs the schools' cooperation...the colleges need the magazine's cooperation just as much. Not because the rankings are accurate or worthwhile, but because they're scared of angering what an Ivy League college president called 'the U.S. News gods,' and falling in the rankings. [Glass, p. 94]

To underscore this last point, in 1995 administrators of Reed College in Oregon decided not to submit ranking data requested by *USNWR*. After several unsuccessful requests, *USNWR* decided to punish Reed by giving the college the lowest possible score in nearly every category. Reed, which had been ranked 18th of all national liberal arts colleges, plummeted to the bottom of its quartile. No other institution had ever fallen so far and so fast.

USNWR has not been insensitive to the numerous criticisms that have been directed towards its college rankings. In addition to tightening up its data collection procedures, *USNWR* modified the criteria weighting scheme and altered its tier structure. Until recently, *USNWR* assigned ranked institutions to one of four tiers, with tier 1 including the "best" colleges and universities in each of six categories (national universities, national liberal arts colleges, and four categories of regional colleges and universities). In terms of academic quality, only institutions in the first tier were ranked in descending order. Institutions in the remaining tiers were listed alphabetically. In its 2004 rankings, *USNWR* reduced from 4 to 3 the number of tiers in the national universities and liberal arts colleges categories. Tiers 1 and 2 were combined into a single tier designated "top schools." It was this modification to the *USNWR* ranking methodology that motivated this updated study.

UNDERGRADUATE RANKING CRITERIA AND WEIGHTS

USNWR bases its undergraduate rankings of national universities and liberal arts colleges on 15 criteria "that education experts have proposed as reliable indicators of academic quality." These criteria fall into seven broad categories: Peer assessment, student selectivity, faculty resources, graduation and retention rates, financial resources, alumni giving, and graduation rate performance. The individual ranking criteria (X_1 through X_{15}) and their weights are summarized in Table 1. The remaining criteria (X_{15} through X_{22}) in Table 1 represent additional information provided by *USNWR* in the ranking tables.

A cursory examination of the *USNWR* ranking criteria suggest the presence of pervasive multicollinearity.[2] An institution's academic reputation (peer assessment), for example, at least partly reflects the SAT/ACT scores of admitted students. The resulting rankings affect tuition and non-tuition based sources of operating revenues, which have an impact on an institution's per-student expenditures, student/faculty ratios, faculty compensation, and so on.

Changes in financial resources have important feed-back effects since higher educational expenditures, lower student/faculty ratios, smaller class sizes, etc. are likely to attract more academically qualified students, which result in higher peer assessments, and so on. The possibility of pervasive multicollinearity between and among the ranking criteria suggests that the assigned weighting scheme may not accurately reflect their actual contributions to *USNWR* tier rankings. This problem is particularly troublesome in the case of peer assessment, which has been assigned a weight of 25%.[3] In 1996, Alma College in Michigan surveyed administrators of 158 national liberal-arts colleges included in the *USNWR* rankings. They found that 84% of the respondents were unfamiliar with many of the schools that they were asked to rank. To make matters worse, one fourth of these respondents just guessed. "That means that reputations were based on what a few people knew about a school, and many of those were just making it up." [Glass, p. 94]

Table 1: *USNWR* Ranking Criteria and Weights

Criteria	Definitions
X_1	Peer assessment--25%. This is a measure of how the institution is regarded by administrators at peer institutions based on a survey conducted in the spring of 2003. (25%)
X_2	Acceptance rate--1.5%. The ratio of the number of students admitted to the number of applicants for the fall 2002 admission. (2.25%)
X_3	High school class standing–top 10%--6%. The proportion of students who graduated in the top 10 % of their high school class. (5.25%)
X_4	Average SAT/ACT scores--7.5%. Average test scores on the SAT or ACT of enrolled students, converted to percentile scores by using the distribution of all test takers. (6%)
X_5	Faculty compensation--7%. Average full-time faculty pay and benefits adjusted for regional differences in cost of living for the 2001-2002 and 2002-2003 academic years. (7%)
X_6	Percent faculty with top terminal degree--3%. Percentage of full-time faculty members with a doctorate or the highest degree possible in their field for the 2002-2003 academic year. (3%)
X_7	Percent full-time faculty--1%. The percentage of the total number of faculty employed on a full-time basis during the 2002-2003 academic year. (1%)
X_8	Student/faculty ratio--1%. Ratio of full-time students to full-time faculty during the fall of 2002. (1%)
X_9	Percent of classes with less than 20 students--6%. The percentage of undergraduate classes (excluding class subsections) with fewer than 20 students enrolled during the fall of 2002. (6%)
X_{10}	Percent of classes with more than 50 students--2%. The percentage of undergraduate classes (excluding class subsections) with 50 students or more during the fall of 2002. (2%)
X_{11}	Average graduation rate--16%. The percentage of freshmen who graduated within a six-year period, averaged over the classes entering between 1993 and 1996. (16%)
X_{12}	Average freshman retention rate--4%. The percentage of first-year freshmen that returned to the same institution the following year, averaged over the period 1998-2001. (4%)
X_{13}	Average educational expenditures per student--10%. The average spending per full-time equivalent student on instruction, research, public service, academic support, student services, institutional support, and operations and maintenance (for public institutions only) during the 2001 and 2002 fiscal years. (10%)
X_{14}	Average alumni giving rate--5%. The average percent of undergraduate alumni who donated money to the institution. These rates were averaged for the 2001 and 2002 academic years. (5%)
X_{15}	Graduation rate performance--5%. The difference between the actual six-year graduation rate for students entering in the fall of 1996 and the predicted graduation rate. (5%)
X_{16}	Graduation and retention rank (derived from factors X_{12} and X_{15}).
X_{17}	Predicted graduation rate. Based upon characteristics (not specified) of the entering class as well as characteristics of the institution.
X_{18}	Faculty resources rank (derived from factors X_5 through X_{10}).
X_{19}	Selectivity rank (derived from factors X_2 through X_4).
X_{20}	Financial resources rank (derived from factor X_{13}).
X_{21}	Alumni giving rank (derived from factor X_{14}).
X_{22}	Dummy variable where 1 designates a public institution and 0 designates a private institution.

Note: Numbers in parentheses represent criteria weights applied to *U.S. News & World Report* 1999 tier rankings.

This paper extends Webster's [2001] analysis of the accuracy of the weighting scheme utilized by *U.S. News & World Report* (*USNWR*) to rank colleges and universities according to "widely accepted indicators of national excellence," which he found to be plagued by severe and pervasive multicollinearity. As in the Webster study, we employ principal component analysis to assess the relative contributions of thirteen criteria used by *USNWR* in 2004 to rank "top schools" in the national university category.

DATA

The data used in this study were obtained from *USNWR 2004 America's Best Colleges* website. *USNWR* categorizes institutions of higher learning as national universities, national liberal arts colleges, regional universities and regional liberal arts colleges according to criteria established by the *Carnegie Foundation for the Advancement of Teaching*. There are 126 institutions included in the 2004 *USNWR* "top schools" tier of national universities.[4]

USNWR reported data on 15 criteria for "top schools" national universities (see Table 1). Unfortunately, of the fifteen ranking criteria used to generate the *USNWR* rankings, data on only 13 criteria are reported on the website. Data on "faculty compensation" (X_5) and "percent faculty with top terminal degree" (X_6) are not reported. Both of these ranking criteria are included in the broader "faculty resources" category. As a partial offset to this omission, *USNWR* provides data on "faculty resources rank," which is constructed using the seven ranking criteria (X_5 through X_{10}) in this category. Of the 126 institutions ranked in the "top schools" national university tier, *USNWR* reports average SAT and ACT scores of the 25th and 75th percentiles.[5] ACT scores were converted to their SAT equivalent utilizing a conversion table published by *Ivy West Educational Services* (www.ivywest.com/acttosat.htm). As in the Webster [2001] study, a single SAT score was obtained by averaging the 25th and 75th percentiles. This paper analyzes 13 of 15 *USNWR* ranking criteria. These 13 criteria account for 90% of the total weight assigned in the ranking methodology.

EMPIRICAL ANALYSIS

Webster [2001] argued that the actual contributions of the ranking criteria examined differ significantly from the explicit USNWR weighting scheme because of the presence of severe and pervasive multicollinearity. Table 2, which summarizes the pair-wise correlation coefficients of the *USNWR* ranking criteria examined in this study, support Webster's conclusions that the weights assigned by *USNWR* to the ranking criteria are not representative of their actual contributions to the rankings. In its college rankings, *USNWR* assigns its greatest weight (25%) to the institution's academic reputation (peer assessment). Not surprisingly, peer assessment (X_1) is very highly correlated with the academic quality of admitted students as measured by average SAT/ACT scores (X_4). SAT/ACT scores are also very highly correlated with acceptance rates (X_2), high-school class standing–top 10% (X_3), student/faculty ratios (X_8), average graduation rates (X_{11}), average retention rates (X_{12}), and average alumni giving rates (X_{14}). Yet, despite its strong correlation with six other ranking criteria, USNWR assigns only a 7.5% weight to average SAT/ACT scores.

It should not be surprising that SAT/ACT scores may be highly correlated with high-school class standing. Also not surprising is that scores on standardized tests are highly correlated with graduation rates. The fact that peer assessment is highly correlated with average SAT/ACT scores, which is highly correlated with average alumni giving, suggests that alumni who are proud of their *alma mater* may believe that their degree provides a competitive advantage in the market place, and are therefore more apt to provide financial assistance after graduation. This linkage helps to explain the strong positive correlation between and among peer assessment, SAT/ACT scores and average educational expenditures per student and student/faculty ratios. The relationship between peer assessment and alumni contributions is important since an institution's ability to provide state-of-the-art educational facilities and hire world-class faculty depends crucially on its non-tuition sources of income. Raising alumni contributions, therefore, may partly depend on higher admission standards.

Peer assessment is also highly correlated with the average freshman retention rate, perhaps because students who believe that the institution's degree is highly valued in the market place are more likely to remain with that institution. There is a possible chain of causality from admission standards to academic quality to retention. Improved retention rates may thus depend on higher admission standards.

Table 2: Correlation matrix of *USNWR* ranking criteria

	X_1	X_2	X_3	X_4	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
X_1	1.00												
X_2	-0.73	1.00											
X_3	0.72	-0.74	1.00										
X_4	0.76	-0.82	0.78	1.00									
X_7	0.33	0.59	0.11	0.04	1.00								
X_8	-0.40	0.56	-0.42	-0.71	0.16	1.00							
X_9	0.35	-0.54	0.44	0.60	-0.17	-0.74	1.00						
X_{10}	0.22	0.11	0.09	-0.17	0.45	0.54	-0.58	1.00					
X_{11}	0.71	-0.76	0.77	0.83	0.05	-0.54	0.52	-0.09	1.00				
X_{12}	0.76	-0.81	0.80	0.84	0.02	-0.49	0.44	0.03	0.90	1.00			
X_{13}	-0.63	0.63	-0.63	-0.67	-0.13	0.62	-0.55	0.08	-0.53	-0.54	1.00		
X_{14}	0.51	-0.64	0.48	0.73	0.06	-0.61	0.50	-0.25	0.64	0.59	-0.42	1.00	
X_{15}	-0.04	0.04	-0.21	-0.26	0.06	0.18	-0.09	0.04	0.19	0.04	0.22	-0.04	1.00

An institution's academic reputation, as summarized in its peer assessment rating, is surprisingly negatively correlated with the institutions acceptance rates. One would ordinarily expect that the higher an institution's academic reputation then the lower its acceptance rate. Paradoxically, the negative correlation coefficient (!0.73) suggests quite the opposite. Webster credited this apparent contradiction to the phenomenon of "self-selection."

"Self-selection refers to the situation in which academically under-qualified students do not apply to tier 1 institutions because, in their judgment, their applications would be rejected. Because of the transaction costs involved in the application process (application fees, time, etc.), the effect of self-selection would be to exert an upward bias on reported acceptance rates for more prestigious institutions. Moreover, as the number of applications of academically under qualified students to institutions of lower prestige increase, acceptance rates at these national universities are biased downwards [2001, p. 239]."

The linkages discussed above suggest that the weight assigned by USNWR to average SAT/ACT scores is too low, while the weights assigned to other ranking criteria, such as peer assessment, average graduation rates and education expenditures may be too high. Webster has argued that an understanding of the relative contribution of the *USNWR* ranking criteria to an institution's overall academic standing is essential to the formulation of the institution's strategic agenda.

Because of severe and pervasive multicollinearity, Webster applied principal component analysis (see, for example, Chatterjee and Price [1977], Hair, Anderson and Tatham [1987], Hotelling [1936], Maddala [1997], and Malinvaud [1997]) to determine whether the arbitrary weighting scheme adopted by *USNWR* is an accurate reflection of the contribution of each of the ranking criteria.[6] The objective of principal component analysis is to derive an alternative linear combination of explanatory variables, called principal components, that have certain desirable statistical properties (See Appendix 1). The estimated coefficients are equal to the eigenvectors of the correlation or covariance matrix. Estimated principal components are then sorted in descending order by eigenvalue, which are equal to the variances of the components. The first principal component should be able to explain variations in the value of the dependent variable better than any other linear combination of explanatory variables. The relative weights of the ranking criteria may be derived from the estimated first principle components.

Table 3 summarizes the eigenvectors of four estimated models. The eigenvector of Model 1 indicates that there are approximately equal loadings on ten of the thirteen ranking criteria examined, which account for about 96 % of the absolute standardized absolute variance, compared with about 88 % in the Webster study. These ten ranking criteria in descending order of their individual contribution to the standardized variance of the first principal components are X_4 (average SAT/ACT scores–11.2%), X_2 (acceptance rate, 10.4%), X_{11} (average graduation rate, 10.3%), X_{12} (average freshman retention rate, 10.3%), X_3 (high school class standing–top 10%, 9.8%), X_1 (peer assessment, 9.4%), X_{13} (average educational expenditures per student, 8.9%), X_{14} (average alumni giving rate, 8.8%), X_8 (student/faculty ratio, 8.7%), and X_9 (percent classes less than 20 students, 8.1%).

An innovation to the Webster study is the estimation of Models 2, 3 and 4, which replaces ranking criteria X_7 through X_{10} with ranking criterion X_{18}. Ranking criteria X_5 through X_{10} constitute the *USNWR* ranking category "faculty resources. Unfortunately, data for X_5 and X_6 were not provided in the *USNWR* website. *USNWR* does provided a ranking of "top schools" according to faculty resources (X_{18}), which were presumably derived from the other ranking criteria in this category. The eigenvector of Model 2 indicate that nine of the ten ranking criteria are approximately equally loaded and account for about 98% of the absolute standardized variance. Once again, average SAT/ACT scores account for the greatest contribution to the first principal component (12.7%).

Table 3: Eigenvectors For The First Principal Components ($N = 126$)

Criteria	Model 1	Model 2	Model 3	Model 4
X_1	0.307	0.332	0.315	0.316
X_2	!0.341	!0.354	!0.345	!0.346
X_3	0.320	0.346	0.336	0.335
X_4	0.366	0.375	0.370	0.370
X_7	0.015			
X_8	!0.285			
X_9	0.265			
X_{10}	!0.071			
X_{11}	0.338	0.351	0.345	0.348
X_{12}	0.337	0.357	0.348	0.349
X_{14}	0.287	0.288	0.286	0.272
X_{15}	!0.045	!0.046	!0.051	
X_{18}		!0.286	!0.285	!0.284
X_{20}	!0.291	!0.296	!0.289	!0.289
X_{22}			!0.206	!0.205

Model 3 modifies Model 2 with the inclusion of a binary dummy variable (X_{22}) where 0 denotes a private institution and 1 denotes a public institution. Of the 126 institutions included in the *USNWR* "top schools" tier, 64 are state-supported universities. The eigenvector of Model 3 indicates that the public school dummy accounts for about 6.9% of the absolute standardized variance. Average SAT/ACT scores remain the most important explanatory variable in the *USNWR* "top schools" ranking.

The final modification removes ranking criterion X_{15} (graduation rate performance) from Model 3. This ranking criterion consistently ranked lowest in its contribution to the absolute standardized variance (less than 2%). The eigenvector for the first principal component for Model 4 is summarized in Table 3. All of the ranking criteria in this model have approximately equal loadings. As before, average SAT/ACT scores is accounts for the largest percentage (12%) of the absolute standardized variance of the complete set of ranking criteria used in the model.

As in Webster's findings, the highest pair-wise correlations exist among these ten ranking criteria. Average SAT/SAT scores, for example, is highly correlated with peer assessment (0.76), acceptance rate (!0.82), high school class standing (0.78), student/faculty ratio (!0.71), average graduation rate (0.83), average freshman retention rate (0.84), and average alumni giving rate (0.73). Although *USNWR* assigns a relatively low weight to SAT/ACT scores (7.5%) its contribution to the rankings, as well as its influence on the remaining criteria, is pervasive.

Table 4 summarizes the estimated eigenvalues of the four models estimated using alternative combinations of the ranking criteria summarized in Table 1. Model 1, for example, summarizes the eigenvalues of the 13 principal components estimated from these data and their proportional contribution to the *USNWR* "top schools" rankings. In principal component analysis, the first principle component is a linear combination of explanatory variables with the greatest variance. The eigenvalues indicate that the first principle component explains 52% of the standardized variance; the second principle component explains another 16%, and so on. Intuitively, the first principal component should provide the best explanation of the *USNWR* rankings. By contrast, the first principal component of Model 4,

which uses only 10 of the ranking criteria listed in Table 1, explain about 66% of the absolute standardized variance. This suggests that Model 4 does a better job at explaining variations in the *USNWR* rankings.

Table 4: PCA Eigenvalues (Proportion Explained)

Principal Component	Model 1	Model 2	Model 3	Model 4
1	6.794 (52.26)	6.380 (63.80)	6.628 (60.26)	6.614 (66.14)
2	2.124 (16.34)	1.243 (12.43)	1.308 (11.89)	1.044 (10.44)
3	1.201 (9.24)	0.625 (6.25)	0.978 (8.89)	0.705 (7.05)
4	0.842 (6.48)	0.555 (5.55)	0.561 (5.10)	0.475 (4.75)
5	0.583 (4.49)	0.411 (4.11)	0.473 (4.30)	0.379 (3.79)
6	0.368 (2.83)	0.267 (2.67)	0.376 (3.43)	0.233 (2.33)
7	0.262 (2.01)	0.228 (2.28)	0.232 (2.11)	0.211 (2.11)
8	0.241 (1.85)	0.171 (1.71)	0.210 (1.91)	0.167 (1.67)
9	0.207 (1.59)	0.085 (0.85)	0.120 (1.09)	0.095 (0.95)
10	0.147 (1.13)	0.035 (0.35)	0.078 (0.71)	0.078 (0.78)
11	0.123 (0.95)		0.035 (0.31)	
12	0.075 (0.58)			
13	0.034 (0.26)			

Note: Eigenvalues represent the column sum of squares for a factor; sometimes referred to as a latent root. It represents the amount of variance accounted for by a factor.

SIMULATIONS AND COMPARISONS

USNWR argues that its "rankings of colleges and universities provide an excellent starting point for families comparing colleges because they offer an opportunity to judge the relative quality of the educational experience at schools according to widely accepted indicators of excellence." The foregoing analysis suggests, however, that the *USNWR* rankings are more appropriately a measure of the academic selectivity of colleges. Does this imply that more selective colleges and universities also provide the highest quality education? The available evidence suggests otherwise. Kuh and Pascarella [2004], for example, have argued persuasively that institutional selectivity is not a good proxy for effective educational practices. This paper is not intended to add to this debate. But, this does raise another question. Are the *USNWR* college rankings an accurate reflection of academic selectivity given the methodological shortcomings in their construction?

Despite their many critics, the general public appears to be enamored with the rankings. They view them as an unbiased third-party assessment of the relative strengths and weakness of colleges and universities. On the other hand, the efforts of administrators to improve their institutions relative position by manipulating selective criteria will continue to fuel the debate over the veracity of the rankings. To whatever ends the *USNWR* rankings are used, however, they ought to be internally consistent in terms of the ranking criteria upon which they are constructed. The underlying structure and character of colleges and universities evolve very gradually. According to one quip, the only thing that moves more slowly than a university is a cemetery. Yet, the annual leapfrogging of institutions suggests

that the *USNWR* ranking methodology is anything but consistent. In this section we will examine the consistency of the 2004 *USNWR* rankings of "top schools."

To examine the consistency of the *USNWR* "top schools" rankings we began by generating alternative college rankings. Principal component regression (PCR) analysis (see Appendix 2) was used to generate the parameter estimates of a linear model of ranking criteria identified in Model 4 of Table 3.[7] The results of this procedure are summarized in Table 5.

Table 5: Principal Component Regression Results With One Principal Component Omitted

Independent variable	Regressioncoefficient	Standard error
Intercept	337.59	
X_1	-14.8392	2.5005
X_2	-0.1646	7.4580E-2
X_3	-3.7271E2	6.2861E-2
X_4	-2.8035E2	1.5843E-2
X_{15}	-0.9662	0.1391
X_{12}	-1.3853	0.2700
X_{14}	3.1660E2	0.1160
X_{18}	0.14	2.3678E2
X_{20}	0.13	2.4436
X_{22}	-7.5641	2.3154
RMSE	62.13	
R^2	0.94	
F-ratio	180.54	

Simulated rankings derived from the PCR model of Table 5 are summarized in Table 6. Column (1) summarizes the *USNWR* "top schools" rankings. Column (3) summarizes the simulated college rankings from the model presented in Table 5. Column 6 is the difference between the *USNWR* and simulated rankings.

How closely do the *USNWR* rankings correspond with the simulated rankings derived from the PCR model summarized in Table 5? A casual review of Table 6 suggests that the lower the *USNWR* ranking (the higher the number) the greater the average variance from the rankings predicted by the PCR model. How significant is this observation? To answer this question we began by calculating the average absolute deviation of the differences between the *USNWR* and PCR model rankings. We found that the first third of the rankings differed by an absolute average of 2.76 places, the second third differed by an absolute average of 5.43 places, and the final third differed by an absolute average of 7.4 places.

**Table 6: *USNWR* "Top Schools" Rankings And Ranking
Simulations From A Principle Component Regression Analysis**

(1) Rank	(2) Institution	(3) Simulations	(4) Difference
1	Harvard University	1	0
1	Princeton University	2	!1
3	Yale University	3	0
4	Massachusetts Institute of Technology	4	0
5	California Institute of Technology	10	!5
5	Duke University	6	!1
5	Stanford University	5	0
5	University of Pennsylvania	7	!2
9	Dartmouth College	12	!3
9	Washington University in St. Louis	19	!10*

11	Columbia University	8	3
11	Northwestern University	11	0
13	University of Chicago	9	4
14	Cornell University	14	0
14	Johns Hopkins University	16	!2
16	Rice University	17	!1
17	Brown University	15	2
18	Emory University	23	!5
19	University of Notre Dame	20	!1
19	Vanderbilt University	25	!6
21	University of California–Berkeley†	21	0
21	University of Virginia	13	8
23	Carnegie-Mellon University	26	!3
23	Georgetown University	24	!1
25	University of Michigan–Ann Arbor†	18	7
26	University of California–Los Angeles†	22	4
27	Tufts University	27	0
28	Wake Forest University	31	!3
29	University of North Carolina–Chapel Hill†	28	1
30	University of Southern California	40	!10*
31	College of William and Mary†	29	!2
32	Brandeis University	34	!2
32	University of California–San Diego†	30	2
32	University of Wisconsin–Madison†	32	0
35	New York University	38	!3
35	University of Rochester	33	2
37	Case Western Reserve University	36	1
37	Georgia Institute of Technology†	43	!6
37	Lehigh University	44	!7
40	Boston College	42	!2
40	University of Illinois–Urbana-Champaign†	35	5
40	Yeshiva University	39	1
43	University of California–Davis†	37	6
44	Tulane University	52	!8
45	University of California–Irvine†	41	4
45	University of California–Santa Barbara†	47	!2
45	University of Washington†	46	!1
48	Pennsylvania State University–University Park†	49	!1
48	Rensselaer Polytechnic University	48	0
48	University of Florida†	45	3
51	George Washington University	60	!9
51	Pepperdine University	57	!6
53	University of Maryland–College Park†	55	!2
53	University of Texas–Austin†	53	0
55	Syracuse University	54	1
55	Worcester Polytechnic University	50	5
57	University of Iowa†	58	!1
58	Purdue University–West Lafayette†	59	!1
58	University of Georgia†	51	7
60	Ohio State University–Columbus†	65	!5
60	Rutgers University-New Brunswick†	56	4
60	University of Miami	75	!15***
63	University of Minnesota–Twin Cities†	68	!5
64	Boston University	62	2
64	Miami University–Oxford†	69	!5

64	University of Connecticut†	61	3
67	Brigham Young University–Provo	84	!17***
67	Indiana University–Bloomington†	76	!9*
67	Texas A&M University–College Station†	66	1
67	University of California–Santa Cruz†	64	3
67	University of Delaware†	72	!5
67	University of Pittsburgh†	73	!6
73	Clark University	83	!10*
73	Michigan State University†	80	!7
73	Southern Methodist University	81	!8
73	University of Missouri–Columbia†	63	10*
73	Virginia Tech†	77	!4
78	Baylor University	94	!16***
78	Clemson University†	92	!14***
78	St. Louis University	79	!1
78	SUNY–Binghamton†	71	7
78	SUNY–College Environmental Science and Forestry†	67	11*
78	University of Colorado–Boulder†	78	0
84	Fordham University	87	!3
84	North Carolina State University–Raleigh†	74	10*
84	University of California–Riverside†	70	14***
87	Illinois Institute of Technology	82	5
87	Iowa State University†	88	!1
87	Stevens Institute of Technology	90	!3
87	University of Denver	89	!2
91	Marquette University	96	!5
91	University of Massachusetts–Amherst†	97	!6
91	University of Tulsa	102	!11*
91	University of Vermont†	86	5
95	Auburn University†	112	!17***
95	University of Kansas†	107	!12*
95	University of New Hampshire†	85	10
95	University of Tennessee†	114	!19‡
99	American University	106	!7
99	Loyola University	95	4
99	Michigan Technological University†	93	6
99	Texas Christian University	115	!16***
99	University of Alabama†	108	!9*
99	University of Arizona†	103	!4
99	University of San Diego	105	!6
99	Washington State University†	101	!2
107	Ohio University†	100	7
107	University of Dayton	104	3
107	University of Kentucky†	121	!14***
107	University of Nebraska–Lincoln†	122	!15***
107	University of the Pacific	91	16***
112	Catholic University of America	99	13**
112	Colorado State University†	110	2
112	Florida State University†	117	!5
112	University of Missouri–Rolla†	98	14***
112	University of South Carolina–Columbia†	124	!12**
117	Howard University	113	4
117	New Jersey Institute of Technology†	118	!1
117	SUNY–Stony Brook†	120	!3

117	University of Oklahoma†	116	1
117	University of San Francisco	109	8
117	University of Utah†	119	!2
123	Clarkson University	111	12**
123	Drexel University	125	!2
123	SUNY–Albany†	126	!3
123	University of Oregon†	123	0

Notes:
1. "†" denotes public institution.
2. "*" indicates statistically significant at the 80% confidence level.
3. "***" indicates statistical significance at the 95% confidence level.
4. "‡" indicates statistical significance at the 99% confidence level.

How significant are these ranking differences? To answer this question we conducted t-tests to determine whether the differences in the *USNWR* and simulated rankings were statistically significantly different from zero. The results of these tests indicate that of the 126 institutions examined, the differences between the *USNWR* rankings and the simulated rankings were statistically significant at the 90% and above confidence levels in 15 cases, or an "error" rate of 12%.[8] Perhaps more interesting is that the number of statistically significant differences appear to increase the lower the rankings. For institutions in the first and second quintiles the percentage error rate is zero. The error rate for institutions in the third quintile is 8%, while the error rates for institutions in the fourth and fifth quintiles are 20% and 31%, respectively. Not only does the error rate increase for lower ranked institutions, but the *USNWR* rankings appear to be negatively biased. Compared with our simulations, at the 90% confidence level *USNWR* overrated 10 of these 15 institutions.[9]

At least in terms of the rankings generated by the PCR model, the *USNWR* rankings become increasing more inconsistent and unpredictable for lower ranked institutions. One possible explanation for this was suggested by the Alma College study cited in Glass [1997], which found that a majority of administrators who were asked to assess peer institutions were unfamiliar with many of the schools that they were asked to rank. It is perhaps not surprising that those administrators (or anyone else, for that matter) should be in greater agreement in their perceptions of Ivy League institutions and less familiar with such institutions as the University of Tulsa or the Illinois Institute of Technology. For less-well-known institutions we would expect wider differences of opinion in peer assessments, and therefore greater variability in the rankings. In some of the most egregious cases of over-ranking, it appears that peer assessments may have as much to do with an institution's prowess on the athletic field rather than its performance in the classroom.

Table 7: Model 4 Principal Components

Criteria	PC 1 (66.1%)	PC 2 (10.4%)	PC 3 (7.1%)	PC 4 (4.8%)	Total (88.4%)
Peer assessment (X_1)	0.316	-0.437	0.023	0.183	0.174
Acceptance rate (X_2)	-0.346	0.095	0.075	-0.073	-0.217
H.S. class standing (X_3)	0.335	-0.207	0.108	-0.343	0.191
SAT/ACT scores (X_4)	0.370	0.062	-0.077	0.049	0.245
Graduation rate (X_{11})	0.348	-0.252	-0.525	-0.096	0.162
Retention rate (X_{12})	0.349	-0.037	-0.305	-0.263	0.193
Expenditures per student (X_{13})	-0.289	-0.162	-0.291	-0.296	-0.243
Alumni giving rate (X_{14})	0.272	0.060	-0.626	-0.126	0.136
Faculty resources rank (X_{18})	-0.284	0.261	-0.360	0.741	-0.151
Public dummy (X_{22})	-0.205	-0.768	-0.038	0.332	-0.197

A clue as to the source of the variability in the rankings may be gleaned from an examination of the first four principal components of Model 4, which explain 88.4% of the *USNWR* rankings. These eigenvectors of the first four

principal components are summarized in Table 7. The first principal component is a rough average of the 10 ranking criteria, with average SAT/ACT scores accounting for the greatest (10%) individual contribution to the absolute standardized variance. Peer assessment ranks sixth in terms of its correlation with the first principal component. By contrast, peer assessment is highly and negatively correlated with the second principal component, accounting for about 19 % of the absolute standardized variance. Only the public dummy variable is more highly correlated with the second principal component, accounting for about 33% of the absolute standardized variance. Whereas the contribution of the public dummy variable in the first principal component is reinforced by its contribution in the second principal component, peer assessment exhibits a significant and sign-changing loading. These results suggest that peer assessment may be the most significant source of inconsistency in the *USNWR* rankings.

SUMMARY AND CONCLUSIONS

This paper expands Webster's [2001] analysis of the accuracy of the weighting scheme utilized by *USNWR* to rank colleges and universities. As in the Webster study, we employed principal component analysis to assess the relative contribution of the *USNWR* ranking criteria in the presence of severe and pervasive multicollinearity. In each of four models estimated, the eigenvectors of the first-principal components confirm Webster's findings that the actual contributions of the ranking criteria differ significantly from the explicit weighting scheme employed in *USNWR* ranking methodology. Contrary to the *USNWR* weighting scheme, the most significant ranking criteria in each case was average SAT/ACT scores of enrolled students, accounting for approximately 11% of the standardized variance of the first principal component. By contrast, *USNWR* assigns a weight of 7.5% to average SAT/ACT scores of enrolled students, which place it fourth in terms of relative importance behind peer assessment (25%), average graduation rate (16%), and average educational expenditures per student (10 %).

This paper also examined the consistency of the *USNWR* rankings of "top schools" by comparing them with the simulated rankings of a principal component regression model. We found that the number of statistically significant differences increases for lower ranked institutions. The absolute differences were statistically significant at the 90 % confidence level in 15 of 126 cases, or an "error" rate of 12%. Moreover, average absolute differences between actual and simulated rankings become larger for lower ranked institutions. Not only did the error rate increase for lower ranked institutions, but the *USNWR* rankings appear to be negatively biased. Compared with our simulations, *USNWR* overrated 10 of 15 institutions at the 90% confidence level. At the 80% confidence level, *USNWR* overrated 17 of 25 institutions.

The results presented in this paper suggest that the *USNWR* rankings become increasing more inconsistent and unpredictable for lower ranked institutions. One source of this inconsistency is peer assessment, which is the only subjective *USNWR* ranking criterion. The reason for this appears to be administrators' lack of familiarity with colleges and universities that they were asked to rate. Moreover, less well known institutions are subject to wider differences of opinion in terms of peer assessments and greater variability in the rankings. This suggest that a possible improvement to the overall consistency of the *USNWR* college rankings would be to lower (or perhaps remove entirely) the relative contribution of peer assessment from the ranking methodology.

ENDNOTES

1. "It sells so well that around *U.S. News'* office it has been dubbed the swimsuit issue." [Glass, p. 94]
2. Multicollinearity in this instance refers to the degree to which changes in the value of one or more of the ranking criteria affect, and are affected by, changes in one or more of the other ranking criteria.
3. *USNWR* measures academic reputation by surveying presidents, provosts, and deans of admissions of peer institutions. These administrators are asked to rate other institutions in their category on a scale of 1 (marginal) to 5 (distinguished).
4. According to *USNWR*, a national university offers a full range of undergraduate majors, and offers masters and doctoral degrees.
5. Of the 126 in the "top schools" tier, 96 reported average SAT scores, while the remaining 30 institutions reported average ACT scores.
6. In geometric terms, principal component analysis is similar to ordinary least squares. In principal component

analysis the *k*-dimensional linear subspace spanned by the first *k* principal components gives the best possible fit to the data points as measured by the sum of the squared perpendicular distances from each data point. Ordinary-least-squares regression analysis, on the other hand, summarized the sum of the squared vertical distances.

7. Principal components regression (PCR) is used for analyzing multiple regression data that suffer from multicollinearity. When multicollinearity is present ordinary least squares (OLS) parameters are unbiased, although large variances may cause these estimates to differ significantly from their true value. PCR reduces standard errors by adding a degree of bias to the regression estimates. In PCA severe multicollinearity is detected as very small eigenvalues. To rid the data of multicollinearity PCR omits principal components associated with small eigenvalues. Regressing the dependent variable against the first and second principal component should eliminate the multicollinearity problem. These results are then transformed back to the original set of explanatory variables to obtain the parameter estimates. Although these estimates will be biased the size of the bias will be more than compensated for by the decrease in the standard errors.

8. At the 80% confidence level this "error" rate increases to 19.8%.

9. At the 80% confidence level, *USNWR* overrated 17 of 25 institutions.

7. REFERENCES

1. Charter, T. Rankled by the Rankings, *ABA Journal*, 84, 1998, pp. 46-53.
2. Chatterjee, S. & B. Price. *Regression Analysis by Example*. New York: John Wiley & Sons, 1997.
3. Crissey, M. Changes in Annual College Guides Fail to Quell Criticisms on their Validity. *The Chronicle of Higher Education*, 44 (2), A67 September 5, 1997.
4. Garigliano, J. U.S. News College Rankings Rankle Critics. *Folio*, 26, 4, 1997, p. 19.
5. Gilley, J. W. Faust Goes to College. *Academe: Bulletin of the AAUP*, 78 (3), May/June, 1992. Reprinted as 'Best College' Lists: The Ranking Game, *Current*, 348, December, 1992, pp. 8-10.
6. Glass, S. The College Ranking Scam. *Rolling Stone*, October 16, 1997, pp. 93-94.
7. Gleick, E. Playing the Numbers. *Time*, 145 (16), 52, April 17, 1995.
8. Graham, H. D. & N. Diamond. Academic Departments and the Rating Game. *The Chronicle of Higher Education*, 45 (41), B6 June 18, 1999.
9. Green, W. H., *Econometric Analysis*, 3rd ed. Upper Saddle River, N.J.: Prentice Hall, 1997.
10. Hair, J. F., R. E. Anderson, & R. L. Tatham. *Multivariate Data Analysis*. New York: Macmillan Publishing Company, 1987.
11. Hotelling, H. Analysis of a Complex of Statistical Variables into Principal Components. *Journal of Educational Psychology*, 24, 1933, pp. 417-441, 498-520.
12. Hotelling, H. Relations between Two Sets of Variables. *Biometrika*, 28 (3-4), 1936, pp. 321-327.
13. Kirk, S. A. & K. Corcoran. School Rankings: Mindless Narcissism or Do They Tell Us Something? *Journal of Social Work Education*, 31 (3), 1995, pp. 408-414.
14. Kuh, G. D., & E. T. Pascarella. What does Institutional Selectivity Tell Us About Educational Quality? *Change*, 36 (5), September/October, 2004, pp. 53-58.
15. Maddala, G. S. *Econometrics*. New York: McGraw-Hill Book Company, 1977.
16. Malinvaud, E. *Statistical Methods in Econometrics*, 2nd ed. Amsterdam, North-Holland, 1997.
17. Marchung, A. Playing the Rankings Game. *Change*, 30 (4), July/August, 1998, pp. 12-16.
18. Morse, R. J. & J. Gilbert. Publishers Perspectives: Magazines. *New Directions for Institutional Research*, 88, Winter, 1995, pp. 91-108.
19. Schatz, M. D. What's Wrong with MBA Ranking Surveys? *Management Research News*, 16 (7), 1993, pp. 15-18.
20. Webster, T. J. A Principal Component Analysis of the *U.S. News & World Report* Tier Rankings of Colleges and Universities. *Economics of Education Review*, 20, 2001, pp. 235-244.

APPENDIX 1–PRINCIPAL COMPONENTS ANALYSIS

Principal components analysis (PCA) is a data-analysis technique that is used to reduce the dimensionality of a large number of interrelated variables, while retaining as much information (variation) as possible. It does this by

calculating an uncorrelated set of variables, called principal components or factors. Each principal component is a linear combination of the selected explanatory variables, with the coefficients equal to the vector of eigenvalues of the correlation or covariance matrix. Eigenvalues represent the column sum of squares for a factor, sometimes referred to as a latent root. Eigenvalues represent the amount of variance accounted for by a factor. These principal components are then ordered in such a way so that the first few of these retain most of the variation in all of the original explanatory variables. The most desirable property of PCA is that the sum of the variances of all of the principle components is equal to the sum of the variances of the original explanatory variables. There are, however, potential problems with this procedure. Although the first principal component has the greatest variance, it need not necessarily be the most highly correlated with the dependent variable. In other words, there is no necessary relationship between the order of the principal components and their degree of correlation with the dependent variable. It is important, therefore, to consider all of the principal components and their proportional explanatory contributions. In addition, the estimated coefficients have no meaningful economic interpretation. Finally, changing the units of measurement of the explanatory variables will change the principal components. To overcome this problem, all of the ranking criteria considered in this study have been standardized to have unit variance.

APPENDIX 2–PRINCIPAL COMPONENTS REGRESSION ANALYSIS

Principal components regression (PCR) is used for analyzing multiple regression data that suffer from multicollinearity. When multicollinearity is present ordinary least squares (OLS) parameters are unbiased, although large variances may cause these estimates to differ significantly from their true value. PCR reduces standard errors by adding a degree of bias to the regression estimates. In PCA severe multicollinearity is detected as very small eigenvalues. To rid the data of multicollinearity PCR omits principal components associated with small eigenvalues. Since the size of the typical eigenvalue of the correlation matrix is unity, PCR eliminates eigenvalues with values considerably smaller than this. For example, if only one small eigenvalue were detected in a problem comprised of three explanatory variables then the third principal component would be eliminated. Regressing the dependent variable against the first and second principal component should eliminate the multicollinearity problem. These results are then transformed back to the original set of explanatory variables to obtain the parameter estimates. Although these estimates will be biased the size of the bias will be more than compensated for by the decrease in the standard errors.