

Modeling Change and Variation in *U.S. News & World Report* College Rankings: What would it really take to be in the Top 20?

Shari L. Gnolek · Vincenzo T. Falciano · Ralph W. Kuncel

Received: 11 December 2012 / Published online: 18 May 2014
© Springer Science+Business Media New York 2014

Abstract University administrators may invest significant time and resources with the goal of improving their *U.S. News & World Report* ranking, but the real impact of these investments is not well known since, as other universities make similar changes, rankings become a moving target. This research removes the mystique of the *U.S. News* ranking process by producing a ranking model that faithfully recreates *U.S. News* outcomes and quantifies the inherent “noise” in the rankings for all nationally ranked universities. The model developed can be a valuable tool to institutional researchers and university leaders by providing detailed insight into the *U.S. News* ranking process. It allows the impact of changes to *U.S. News* subfactors to be studied when variation between universities and within subfactors is present. Numerous simulations were run using this model to understand the effect of each subfactor individually and to determine the amount of change that would be required for a university to improve its rank or move into the top 20. Results show that for a university ranked in the mid-30 s it would take a significant amount of additional resources, directed in a very focused way, to become a top-ranked national university, and that rank changes of up to ± 4 points should be considered “noise”. These results can serve as a basis for frank discussions within a university about the likelihood of significant changes in rank and provide valuable insight when formulating strategic goals.

Keywords U.S. News Rankings · Modeling · Variation · National Universities

S. L. Gnolek (✉)
Scannell & Kurz, 71-B Monroe Ave., Pittsford, NY 14534, USA
e-mail: gnolek@scannellkurz.com

V. T. Falciano
University of Rochester, 500 Joseph C. Wilson Blvd., PO Box 270021, Rochester, NY 14627, USA
e-mail: vfalciano@admin.rochester.edu

R. W. Kuncel
University of Redlands, 1200 East Colton Avenue, PO Box 3080, Redlands, CA 92373, USA
e-mail: ralph_kuncel@redlands.edu

Introduction

The college rankings calculated by *U.S. News & World Report* have an impact on the decisions of students and university leaders and are the most widely circulated scorecard of university performance. One can argue their merits, but two things are clear: rankings fill a strong consumer demand, and they are not going away any time soon. University administrators often criticize the choice of criteria and their assigned weights as being subjective and biased, but they are compelled to spend energy competing in the academic marketplace, often putting in place competitive strategies to try to enroll highly qualified students, raise funds, and recruit faculty in order to improve their overall ranking (Grewel et al. 2008).

In addition, researchers have accurately pointed out that while “changes” in rank generate revenue for ranking publishers, rankings only complicate the lives of university administrators and those in charge of enrollment management (Ehrenberg 2003a). Others have shown that the *U.S. News* ranking process magnifies insignificant, random changes in the underlying subfactors creating small differences in the overall rank (Sauder and Lancaster 2006). Yet when these changes are reported, consumers, trustees, alumni, and university leaders may not have a clear picture of how to evaluate whether the difference is “noise” or a meaningful upward (or downward) change.

While many universities have set strategic goals of improving their rankings and have communicated that objective to board members, faculty, and alumni, developing a clear understanding of what would be required to move up in rank can be a perplexing problem. Rankings are “sticky” and difficult to change (Grewel et al. 2008), and remarkably stable, especially within the top tier (Bowman and Bastedo 2009). Each of the highly correlated subfactors has its own inherent variability, and each university has a unique starting point as it attempts to improve in the rankings. It has been shown that jumping into the top quartile has more of an impact than a shuffling of ranks within the group (Meredith 2004), but research has not been conducted to empirically determine how a university could move into the top 20, or if that is even possible, when not holding comparative institutions constant.

Related Literature

The Influence of Rankings

U.S. News college rankings play a role in student decision making. Parents and students sometimes use them as a starting point to help them filter through information and narrow options when beginning the college search process, and there is evidence that shows they consider them when making decisions about where to enroll (Griffith and Rask 2007). Results from the Fall 2011 Cooperative Institutional Research Program freshman survey of the Higher Education Research Institute showed nearly one quarter of first-time, full-time freshman selected “rankings in national magazines” as “very important” in deciding to go to a particular college. Moreover, rankings shape students’ early perceptions of college quality, creating a normalized view of what is “best,” and affecting students’ perception of their probability of acceptance (Bowman and Bastedo 2009).

Given their influence on admissions outcomes, resource allocation, and strategic planning, *U.S. News* rankings also have a tangible impact on universities. Rankings have been shown to be especially important to private institutions, regardless of their ranking

tier, because universities alter their behavior in response to them (Ehrenberg 2003a). Institutions go to great lengths to manage the perception that comes from a drop in rankings and to correct errors in their submitted *U.S. News* inputs if they occur. Administrators often call for improving rank through highly weighted subfactors, even though the subfactors may not improve actual quality or recruitment (Ehrenberg 2003a). Some universities will actively solicit applications from less qualified students in order to increase the number of applications or reject top students if it is believed they want to go to better schools. Early decisions processes are put in place or SAT scores may become optional because of anticipated effects on rankings (Ehrenberg 2003b; Meredith 2004).

Other behaviors point to the importance of rankings as well. Because tuition costs and instructional expenses signal quality and affect admissions outcomes, some colleges have actually increased their tuition as a way to try to become “elite” (Bowman and Bastedo 2009). Universities will solicit nominal contributions from alumni to boost their percentage of alumni giving, increase scholarship money but direct it toward merit aid (versus need-based aid), create honors colleges to attract more highly qualified students, or send expensive publications to key administrators in order to improve their peer assessment scores (Ehrenberg 2003b; Farrell and Van Der Werf 2007; Sauder and Lancaster 2006).

Rankings are also factored into strategic plans. As one example, the 2002 ten-year strategic plan for Baylor University outlined goals to move their *U.S. News* ranking from tier 2 into the top 100. As of 2007, Baylor had spent \$200 million on related improvements (Farrell & Van Der Werf 2007) and was ranked 75th in the national university category in 2012, the end of the ten-year period. In another example highlighting the broadening impact of rankings, a performance bonus for the president of Arizona State University was tied to the university’s *U.S. News* ranking (Jaschik 2007). Almost every university examines its *U.S. News* submitted data, looking for ways to make legitimate adjustments (Yablon 2001). These behaviors are not unwarranted. Other factors being held constant, when a university improves its rank it can expect a positive benefit the next year as measured by more applications, higher enrollment, and improved SAT scores, all with less financial aid (Monks and Enrenberg 1999; Ehrenberg 2003b). In fact, the draw of a university’s ranking can lure high-income, high-ability students more than the effect of merit aid (Griffith and Rask 2007), having a direct impact on a university’s financial position in measures such as net tuition revenue and tuition discount rate.

Problems Associated With Rankings

The focus on wealth, fame, and exclusivity that comes from placing emphasis on *U.S. News* rankings creates very real issues and highlights some of the inherent problems with the rankings themselves. As institutional behaviors shift when responding to the perceived concept of quality embedded within the ranking process, it is predicted that schools will become more like what is being measured, which will not always align with policy goals particularly in the areas of equity and diversity (Sponsler 2009). Localized competition among similarly ranked universities may ultimately weaken institutions (Grewel et al. 2008) and cause many of them to become less distinguishable from the very universities they compete against (Schmidt 2008). In fact, some institutions have changed their very nature in order to meet strategic goals related to *U.S. News* rankings (Farrell and Van Der Werf 2007).

Further, the ranking methodology employed by *U.S. News* has the appearance of being volatile, creating false precision that can have significant consequences on behavior. Early research suggested that aggregating noisy and reversible subfactors was the most likely

cause for *U.S. News* ranking changes (Dichev 2001). Later work determined that for 68 % of law schools in the top 50 graduate schools, which uses a different methodology that emphasizes reputation scores, a change in rank mainly comes from small changes in the distributions of underlying subfactors (Sauder and Lancaster 2006). Similarly, it has been shown that *U.S. News* rankings are transitory and reversible, with 70–80 % of changes reverting over the following two ranking periods, and “real” change accounting for only about 10 % of the differences (Dichev 2001). The weighting scheme within the methodology of the *U.S. News* ranking process has also been questioned (Clarke 2004). Overall ranks are computed despite a lack of research on the indicators and their weights (Machung 1998), and consultants have concluded that the weights assigned by *U.S. News* lack any empirical or theoretical basis (Farrell and Van Der Werf 2007). Even after working closely with *U.S. News* contacts, attempts to recreate each university’s predicted graduation rate were unsuccessful (Betsinger 2009), so institutions are judged against a graduation standard that is not well understood and has a weight of 7.5 %, more than ten other subfactors.

With stakes high, universities have an incentive to publish inaccurate or misleading data or engage in manipulative behavior. Colleges have been known to “game” the system by falsifying SAT data (Adams 2012) or by rating programs other than their own as below average when filling out peer assessment forms (Lederman 2009). Even more troubling, it has been shown that once published, college rankings have a significant impact on any future peer assessments, independent of college quality and performance (Bastedo and Bowman 2010). A 2011 report by the National Association for College Admission Counseling described how institutions feel pressured by the *U.S. News* ranking process to invest in strategies to maintain or improve their rankings, which leads members to manipulate numbers or adopt targeted strategies merely to improve their rankings (National Association for College Admission Counseling 2011). This behavior creates an environment in which emphasis is placed on making small adjustments to the *U.S. News* inputs, but with limited ability to predict what the final impact might be. Even Robert Morse, Director of Data Research for *U.S. News*, was “highly suspicious that some fudging occurred,” although he didn’t believe it had a large impact mainly because a university can’t manipulate its own reputational data, which then accounted for 25 % of the overall score (Webster 1992).

Ranking Longevity

The *U.S. News* rankings have become a fact of life in the higher education marketplace. At one point they were regarded as the least problematic of the rankings (Webster 1992), but are still the most commercially successful, and are recognized as a performance scorecard by students and administrators alike. The published form of the ranking creates a broadly understood stratification of who is “elite,” which signals legitimacy to policy makers and the public (Bowman and Bastedo 2009). With the increased mobility of students, and the fact that rankings are promoted by colleges and have an international reach, policy makers consider them a durable fixture in the landscape of higher education (Sponsler 2009). Their influence has grown, and the most desirable prospective college applicants are the ones who use them (McDonough et al. 1998). Students attend college not only to prepare themselves broadly, but also for economic well-being, and are thereby taking action to increase their income potential. By one estimate, 50–55 % of students sampled chose to enroll in the highest ranked school that they were admitted to, and this choice was most predominate with high-ability students (Griffith and Rask 2007), so it may be in a university’s best interest to participate in the ranking process.

Prior *U.S. News* Modeling Research

A great deal of research has been done to understand what influences the *U.S. News* rankings and how much movement could be expected. In an early approach, based on the jackknife method coupled with linear regression, researchers examined only the top-50 national universities (the original number in the ‘top’ tier) and recommended that schools be placed into quality bands based on the conclusion that much of the change in rank was “noise” versus a real change in quality (Clarke 2002). In another approach, investigators used principal components analysis to examine the relative contributions of the ranking subfactors for those national universities in the top tier that had reported SAT scores. They found that the actual contribution of each subfactor differed substantially from the weights because of correlation among the variables. These researchers argued that “an accurate understanding of the actual contribution of the ranking criteria to an institution’s overall academic standing is an essential element in the formulation of a university’s strategic agenda” (Webster 2001). In later work, other researchers developed a logit model to address the localized nature of ranking competition, again within the top-50 national universities. They concluded that rankings are “sticky” and that change will happen in incremental steps. Their model used only the *U.S. News* subranks as input and predicted that for a university ranked at 15, 90 % of its change will be within 4 points of its current rank (Grewel et al. 2008). Further, they suggested that more research be done to understand the competitive dynamics that come from ranking universities.

In the analysis presented here, we acted on the suggestions of prior researchers and addressed the gaps in earlier findings by empirically determining how *any* nationally ranked university could move its *U.S. News* ranking, or if it is even possible to do so, when competition exists and rankings of comparative institutions are not held constant. Further, we clearly defined a meaningful change in rank and determined what should be considered “noise” along the entire national university category. Specifically, we addressed these research questions:

1. Given that the *U.S. News* ranking subfactor inputs are highly correlated and subject to their own variability, what specific changes would be required for a university to become more highly ranked within the national university category?
2. Across all 200 schools within the top tier of the national university category, what amount of ranking change should be considered a meaningful difference, and what should be attributed to “noise”?

Model Development

In order to answer these two questions, a model was created that uses the same methodology described by *U.S. News* (Morse and Flanigan 2011) and predominately the same data, supplementing inputs with publicly available data. This model is accurate to within the noise threshold determined by this research and meets the statistical assumptions required of a viable model with errors that are normally distributed, homogeneous, and independently distributed. The Integrated Postsecondary Education Data System (IPEDS) shows that, of the first-time, full-time freshmen enrolled in more than 2,000 primarily baccalaureate (or above) degree-granting U.S. higher education institutions, 45 % are within schools covered by this research.

Input Data

This model was built based on the 2012 *U.S. News* rankings of schools within the national universities category (released in 2011 from mostly fall 2010 data). It included all 268 schools, including those given a Tier-2 ranking. One school (ranked #150 in 2012) was removed because of missing data. The model used published *U.S. News* subfactor data when available. For the thirteen subfactors italicized in Table 1, data were accessible through *U.S. News* and were used exactly as described by the *U.S. News* ranking process. Table 1 also lists the weights assigned by *U.S. News* to each subfactor.

Data for the remaining three subfactors that are not published by *U.S. News* (SAT/ACT scores, faculty compensation, and financial resources per student) were collected from IPEDS, and the subfactors were estimated according to *U.S. News* protocols. *U.S. News* uses the percentile value of the actual reported SAT and ACT *average*, weighted by the percent of students submitting results for each test. Without access to the submitted average scores, the percentile value of the *median* SAT and ACT scores were estimated using the 25th and 75th percentiles from either IPEDS for the SAT scores or *U.S. News* data for the ACT scores. These median values were then weighted, as per *U.S. News* procedure, based on the percent submitting SAT and ACT scores reported to IPEDS.

In order to estimate the financial resources per student subfactor, expenses for research, public service, instruction, academic support, student service, institutional support, and (for public universities only) operations/maintenance were obtained from IPEDS. A logarithmic transformation was applied to the two-year average per student, and this value was weighted based on full-time enrollment.

Finally, *U.S. News* uses a two-year average of faculty compensation in their ranking process and adjusts each year based on the Runzheimer cost of living index. To create a reasonable model using publicly available data, the American Chamber of Commerce Researchers Association (ACCRA) cost of living index compiled by the Council for Community and Economic Research was used as a starting point when adjusting faculty compensation. The Runzheimer index was available to the authors for all universities with membership in the Association of American Universities Data Exchange (AAUDE), so this index was also taken into consideration for AAUDE universities, and adjustments were made to align the model predictions and *U.S. News* rankings.

Significant Change versus “Noise”

The first challenge in creating a valid model and interpreting the significance of changes in a *U.S. News* ranking lies in being able to address the secondary research question and accurately separate a real signal from the noise. To establish a noise threshold, ten years of *U.S. News* rankings data were examined for universities that were ranked in the top 200 for the national university category in 2012, and that had been in the top 200 for five or more of the last 10 years ($N = 130$). This analysis showed that the variance in a university's ranking is a function of where a school is initially ranked. Given the same sample sizes, better-ranked schools experience less variation, whereas greater fluctuations in lower-ranked schools are more likely. The difference between a school's rank in any year and its own historical average gives a clear indication of what type of change could be expected. A confidence interval around that difference quantifies where a meaningful difference is likely to fall.

Figure 1 shows the 95 % confidence bounds on the average difference for each university during the ten-year time period. Any differences within these bounds would not be

Table 1 U.S. news ranking categories, subfactors and weights

Category	Subfactor name	US News Weighting (%)
Undergraduate academic reputation	<i>Peer assessment survey</i>	15.0
	<i>High school counselor's rating</i>	7.5
Student selectivity	<i>Acceptance rate</i>	1.5
	<i>High school standing in top 10 %</i>	6.0
	Critical reading and math portions of the SAT and the composite ACT score	7.5
Faculty resources	Faculty compensation	7.0
	<i>Percent faculty with top terminal degree</i>	3.0
	<i>Percent faculty that is full-time</i>	1.0
	<i>Student/faculty ratio</i>	1.0
	<i>Class size under 20</i>	6.0
	<i>Class size over 50</i>	2.0
Graduation and retention rate	<i>Average graduation rate</i>	16.0
	<i>Average freshman retention rate</i>	4.0
Financial resources	Financial resources per student	10.0
Alumni giving	<i>Average alumni giving rate</i>	5.0
Graduation rate performance	<i>Graduation rate performance</i>	7.5

considered statistically significant. For most universities ranked in the top 40, differences in rank within about ± 2 should be considered noise. For example, a university ranked at 30 could be 95 % confident that its rank will fall between 28 and 32, and only when the rank moves beyond those levels can a statistically significant change be claimed. For universities with ranks beyond the top 40, differences of ± 4 (or even greater) should be considered noise. A few universities have much larger confidence intervals of ± 10 or more. In these specific cases (Arizona State University, Brigham Young University, Drexel University, and Northeastern University) the variability can be correlated to institutional changes during this time period, which resulted in supplying notably different inputs to *U.S. News*.

Model Development

A predictive model was built by applying the same subfactor weights and Z-score methodology used by the *U.S. News* ranking process (Morse et al. 2011). First, the subfactor value for each school was subtracted from the average subfactor value across all schools, then divided by the standard deviation of the subfactor across all schools, resulting in a Z-score for each subfactor. This process was repeated for each subfactor. Next, each subfactor was multiplied by the weight assigned by *U.S. News* (Table 1) and summed to calculate an overall score. This overall score was divided by the highest score across all schools and multiplied by 100. The descending rank of this transformed score, rounded to the nearest whole number, was the final rank assigned by the model.

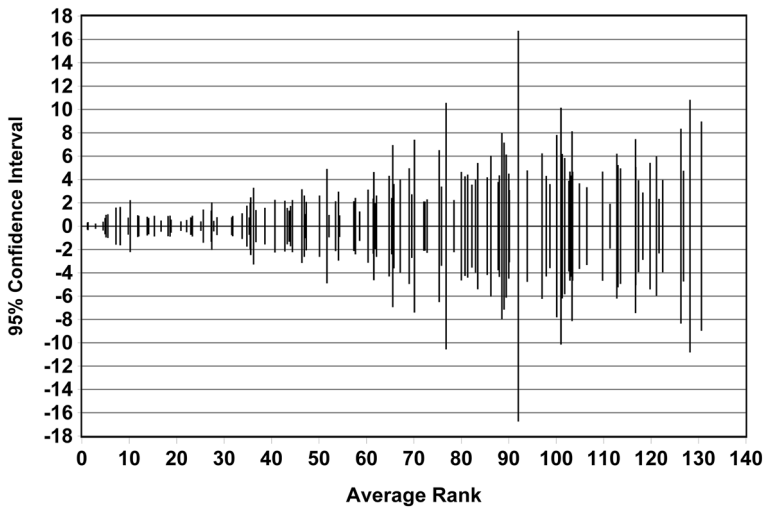


Fig. 1 Confidence intervals of the average difference in rank depend on a university's average placement within the ranking over 10 years

Model Properties

The resulting model credibly reproduces the *U.S. News* ranking results for 2012. It provides ranking predictions within the level accepted as statistical noise across all 268 schools falling into the national universities category. Plotting the rank predicted by this model against the actual 2012 *U.S. News* rank graphically illustrates the precision to which estimates can be made (Fig. 2). This result is not unexpected given that we used the same process as *U.S. News* with largely, but not completely, the same input data. Faculty compensation and SAT/ACT score, the two subfactors that had to be estimated, are the most likely sources of any difference between the model and the actual 2012 *U.S. News* rank.

The error in this model was calculated by subtracting the rank predicted by the model from the actual 2012 rank assigned by *U.S. News*. All of the errors are within ± 4 points, and in 21.39 % of the cases the predicted rank exactly matched the *U.S. News* rank (Fig. 3).

Model Limitations

Undergraduate Academic Reputation Score

One of the largest components in the *U.S. News* rankings is the undergraduate academic reputation score, which counts for 22.5 % of the rank. In 2011 the format of this reputation score was changed to include high school counselor ratings. Prior to 2011 the rating was based strictly on a peer assessment and was rated on a 1 (marginal) to 5 (distinguished) scale. When examining the variability in the overall *U.S. News* rank it is important to point out the consistently low variance in the reputation score over the eight-year time period as seen in Fig. 4. The variability is nearly constant across all overall ranks. Whether a university is ranked high or low, the amount of change expected in its peer assessment

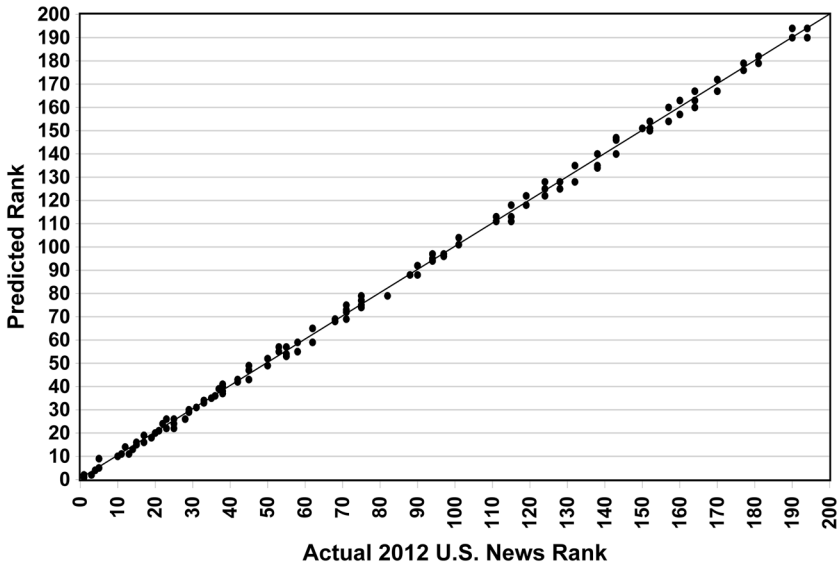


Fig. 2 The rank from the model accurately predicts the *U.S. News* published rank across the entire range of ranked national universities

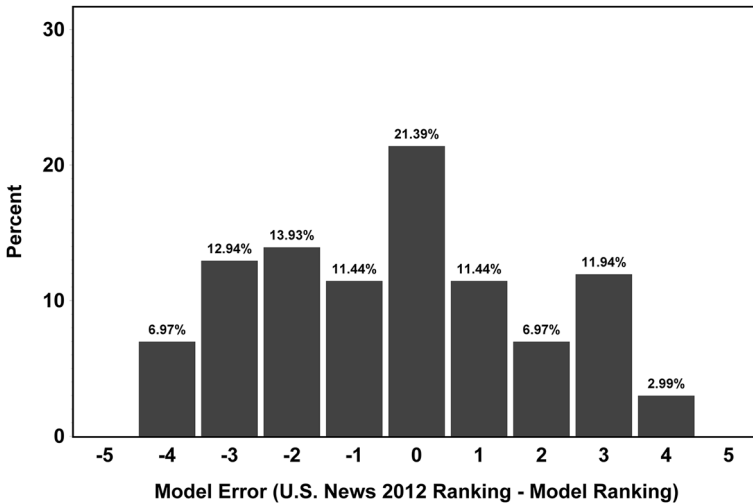


Fig. 3 Examining the distribution of the error in the model shows that all predicted ranks are within four points of the *U.S. News* published rank

score is small (a standard deviation of less than 0.10, which translates to a peer assessment score change of less than 0.10 on the 5 point scale). In fact, four highly ranked universities experienced no change at all in their undergraduate academic reputation score during the entire eight-year time period (Harvard University, Massachusetts Institute of Technology, Stanford University, and the University of Rochester). Based on this analysis, some

simulations using the model were done without perturbing the reputation score, as described in more detail later.

Graduation Rate Performance

Graduation rate performance, the difference between a school's actual graduation rate and what *U.S. News* predicts it should be, was restricted during simulations to be randomly sampled around a school's current level. Changing this subfactor, without being able to accurately recreate the *U.S. News* predicted graduation rate, could have given an undue advantage (or handicap) to each institution. At the time of this research, details were not available as to how the *U.S. News* graduation rate model was created, and it was clearly not always a good predictor of the actual graduation rate. While the differences between the *U.S. News* predicted graduation rates and actual graduation rates are centered on zero across all top 200 schools, implying that the model has no inherent bias across the entire group, the predicted graduation rate can be off by as much as 20 % for any individual school (Fig. 5). In an independent examination of the robustness of the graduation rate performance indicator, Porter (2000) concluded that it was impossible to use the *U.S. News* method to claim that any university was under or over-performing in a meaningful way. A researcher attempting to accurately recreate the *U.S. News* predicted graduation rate was unsuccessful, even with close guidance from *U.S. News* (Betsinger 2009). After careful consideration, we did not attempt to recreate the *U.S. News* graduation rate model and develop a new predicted graduation rate estimate as the underlying subfactors changed. Instead, we assumed the amount of graduation rate performance variation for each school continued to remain around its current level.

Simulations Using the Model

The robust model created for predicting *U.S. News* rankings serves as a valuable tool to estimate results of the *U.S. News* ranking process and to communicate potential outcomes. For example, this model can be used to determine what the rankings might be if underlying subfactors didn't change for any of the universities but instead just varied within the ranges typical for each school. It can also be used to determine the impact of potential changes in one subfactor on a university's final rank. More importantly, it provides an avenue for addressing the first research question; what specific changes would be required for a university to become more highly ranked within the national university category when ranking subfactors are highly correlated and subject to their own variability? Each of these three scenarios is examined through simulations using this model.

Simulation Assumptions and Limits

For all of the simulations performed, the subfactor values for each institution were randomly sampled using the inherent variability from a population that was centered on its own average and varied within ± 3 standard deviations. The mean and standard deviation were based on the last five years of *U.S. News* submission data for each school, with a few exceptions.

Only one year of *U.S. News* subfactor data were available for all 268 schools for the median SAT/ACT, percent of faculty with top terminal degree, and financial resources per student subfactors. In order to provide a reasonable estimate of the variability for these

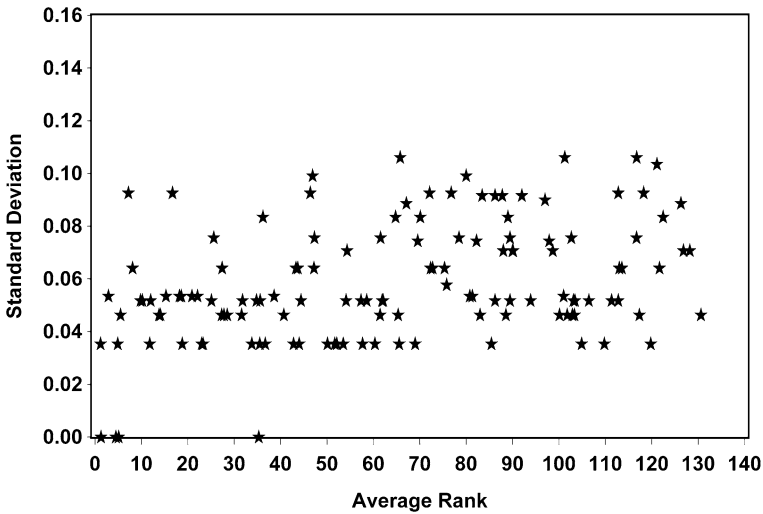


Fig. 4 The standard deviation of the (pre-2011) *U.S. News* reputation score shows there is very little change over this eight-year time period in a university’s reputation score, regardless of its rank position

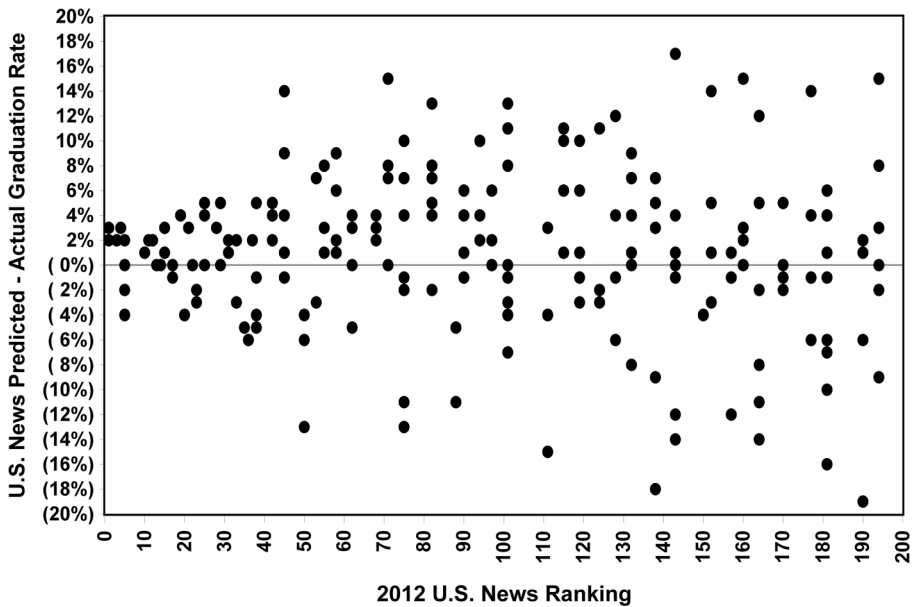


Fig. 5 There are large differences between the graduation rate predicted by *U.S. News* and a university’s actual graduation rate (graduation rate performance) across the entire range of top-200 schools

subfactors a standard deviation was assumed based on data from the University of Rochester, a mid-ranked school where historical information was accessible. Additionally, only one year of input data was available for the faculty compensation subfactor. Rather than assume a fixed value for the standard deviation for this subfactor, which could

potentially be inaccurate at both ends of the faculty compensation spectrum, the standard deviation for this subfactor was assumed to be 2 % of each school's average faculty compensation value.

The new format for the two components of the academic reputation score (peer assessment and high school counselor assessment) had only been in use for two years. For these subfactors, the 2012 value was used without any perturbation, because any estimate of the standard deviation would have been inflated with only two samples. As such, this simulation represents the effect of changes in all other subfactors when the undergraduate academic reputation score remains at 2012 levels. Finally, subfactors were not allowed to go beyond natural (or reasonable) limits. The limits set for the simulation process (if any) are shown in Table 2.

Findings

Predicting Rankings When Universities Vary Just Within Typical Ranges

Using this model, a set of 2,000 Monte Carlo simulations were run to simulate the results of the *U.S. News* ranking process on national universities in the top 200 if all universities wavered just around their own historical averages, and did not improve (or decline) in any of the subfactors. A similar result to that described earlier is found; universities ranked in the top 40 have less variation than those below 40. The ranking variation is slightly higher, because the subfactors are being actively varied, but the trends remain the same (Fig. 6).

Predicting the Impact of Potential Changes in a Subfactor on a University's Final Rank

Oftentimes university leaders will want to understand the impact of potential changes in a single subfactor on the overall rank. While the subfactors are not always independent of each other, it is possible to estimate the changes that would have been required for a single subfactor to move the overall rank in 2012. Since it has been shown the variation in rank is related to where a school falls in the rankings, and that there are differences in the impact each subfactor will have based on its own variation and assigned weight, the model should be run separately for the specific subfactor values for each university to get the most accurate estimates of the change required for that university.

For illustrative purposes, the effect of each subfactor for the University of Rochester, a representative school ranked in the mid-30 s, is shown in Table 3. This table shows the original subfactor value submitted to *U.S. News*, as well as what the model predicts each value would need to have been in order for that single subfactor to drive up the overall 2012 rank *one point* (assuming all other schools stayed at 2012 levels). The difference between the original value submitted and the predicted value required for a one-point change in rank is shown in the last column of the table, in the units for each subfactor. For example, the average graduation rate, which has a weight of 16 % in the *U.S. News* rankings, would have needed to go from 84 to 86 %, a difference of 2 %, in order to improve the overall rank one point. For a subfactor with a smaller weight, such as acceptance rate, a much larger movement would be required: in this case, from 38 to 8 % (or a more than four-fold improvement in acceptance rate). When a subfactor had a small weight and/or not much difference among institutional values, it was not possible to move the subfactor enough to increase the rank by even one point. Given that the subfactors are *not independent* of each other, it would not be accurate to add up all of these changes and

Table 2 Limits set on the subfactors during Monte Carlo simulation

Criteria	Simulation limits
Peer assessment	2012 value
High school counselor assessment	2012 value
Acceptance rate	5–98 %
High school standing in top 10 %	5–99 %
Median SAT percentile (score)	≤99th percentile
Median ACT percentile (score)	≤99th percentile
Faculty compensation	No limit set
Percent faculty with top terminal degree	≤99 %
Percent of faculty that are full-time	≤99 %
Student/faculty ratio	≥3:1
Class size under 20	≤99 %
Class size over 50	>0 %
Average graduation rate	≤99 %
Average retention rate	≤99 %
Financial resources per student	No limit set
Average alumni giving rate	≥Minimum in last 5 years
Graduation rate performance	No limit set

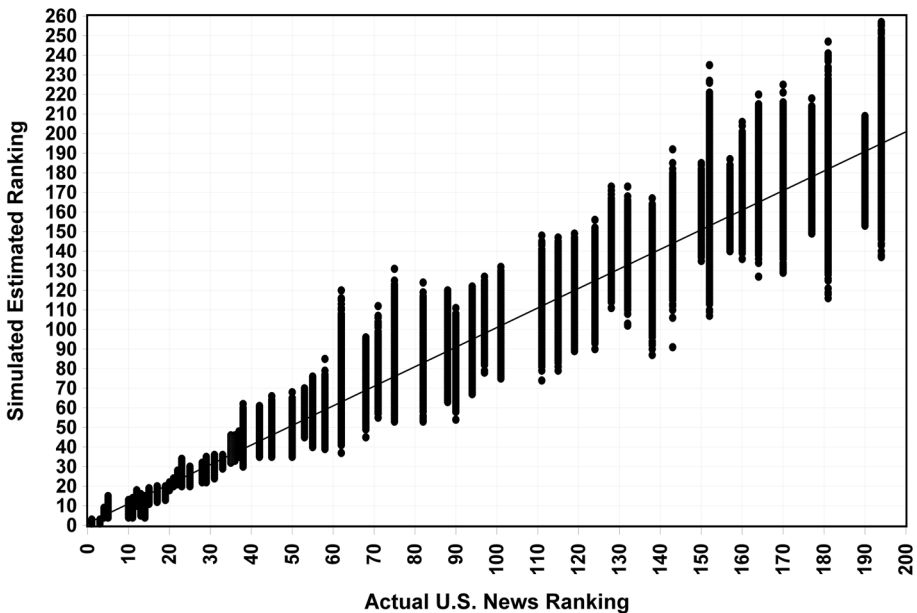


Fig. 6 Monte Carlo simulation using the inherent subfactor variability confirmed that universities ranked in the top 40 would be predicted to have substantially less variation than those ranked below 40

expect that a 17-point change in rank could be reached. Instead, any combined changes should be evaluated together using the model as is shown in the next section. Table 3 could be generated for any university using this model and would be a useful tool for illustrating

the overall impact of *U.S. News* subfactors individually. The table provides details of the “net effect” of each subfactor when its weight and variability are taken into consideration.

In this simulation, each subfactor was evaluated and changed *independently* of the other subfactors. Correlations between subfactors would not have an impact, so were not evaluated.

To ascertain whether the changes predicted by the model were reasonable, the inherent variation in the underlying subfactors was examined and similarly weighted subfactors were compared. Probability density curves for two selected subfactors can be used to illustrate this approach. The graphs in Fig. 7 show the distribution of the percentage of full-time faculty (left) and acceptance rate (right), separated into ranking groups of ten universities (only the top 70 schools are shown for simplicity). The full-time faculty percent subfactor has a weight of only 1.0 % in the ranking process. Because of the small differences between ranking groups, and the low weighting, it is clear that creating movement in the overall rank with this subfactor would be difficult. Given this, it appears that the model prediction was reasonable; changes to this subfactor alone will not move the overall rank even one point. Acceptance rate also has a similar weighting (1.5 %). However, with this subfactor there is a larger difference between the acceptance rates for each ranking group as well as larger variation within a group, so there is greater opportunity for change with this subfactor than the similarly weighted full-time faculty subfactor. In this example, the model predicted that moving the acceptance rate 30 %, from 38 to 8 %, would move the overall rank one point. Comparing the probability distribution functions for the 31–40 ranking group (where a mid-30 s university would fall) to that of the groups above it shows that the model prediction for acceptance rate also appears reasonable. Each subfactor was examined in this way, but the others are not included here for brevity.

Predicting What it Would Take to Move into the Top 20

Finally, simulations were performed to address the primary research question and determine what specific changes would be required for such a university to become more highly ranked within the national university category. Specifically, for a university ranked in the mid-30 s, what would be required to achieve a top-20 *U.S. News* ranking? This model was used to determine the answer to that explicit question but could also be used to address movement in rank along any point in the range. There are many possible ways to move, each involving a combination of subfactors, depending on where a university initially starts in the overall ranking (and with each subfactor).

Another set of 2,000 Monte Carlo simulations were run. In this case, the subfactor values for each school were again randomly sampled and varied within ± 3 standard deviations, but the example institution was given an improved starting point for each of the subfactors. Universities have some level of direct control over each of the subfactors, with the exception of reputation score, which is derived from the peer assessment and high school counselor ratings. With the reputation score carrying a combined weight of 22.5 % and the fact that this score has been shown to change very little over time, it would be difficult, if not impossible, to move into the top 20 without achieving an undergraduate reputation score equal to that of universities at that level. As such, this simulation was run using two different sets of conditions. In one case, it was assumed that the reputation score for this school would move into the range typical of a university that is ranked between 18 and 22 for the peer assessment score (4.2) and the high school counselor score (4.5). In the other case, the reputation score was left at the current level. All other simulation limits listed in Table 2 were still applied.

Table 3 Subfactor changes required to move up one overall rank for a university ranked in the mid-30 s

Criteria	US News 2012 submitted value	What this <i>single subfactor</i> would have had to be, to move up <i>one</i> overall rank in 2012	Movement required for a one-point change in rank
Peer Assessment	(rated 3.4)	3.6	0.2
High School Counselor Assessment	(rated 3.9)	4.0	0.1
Acceptance Rate	38 %	8 %	–30 %
High School Standing in Top 10 %	79 %	87 %	8 %
Median SAT percentile (score)	90 % (~1,300)	96 % (~1,400)	6 % (+100)
Median ACT percentile (score)	96 % (~30)	Not possible	
Faculty Compensation	\$140,793	\$147,293	\$6,500
Percent faculty with top terminal degree	94 %	99 %	5 %
Percent faculty that is full-time	86 %	Not possible	
Student/faculty ratio	10	Not possible	
Class size under 20	69 %	74 %	5 %
Class size over 50	11 %	5.5 %	–5.5 %
Average graduation rate	84 %	86 %	2 %
Average retention rate	96 %	Not possible	
Financial resources per student	\$73,923	\$81,323	\$7,400
Average alumni giving rate	17 %	20 %	3 %
Graduation rate performance	–5.0 %	–3.0 %	2 %

In this simulation, changes were made to all of the subfactors individually, but were evaluated in the context of the entire set of changes. The Monte Carlo simulation process was used to estimate a value for each subfactor by sampling from the distribution of each subfactor independent of the distributions of other subfactors, so conditions that may appear unlikely (i.e., increasing SAT/ACT score and decreasing graduation rate) would have been created by the process. These conditions would not have led to an improvement in rank and were not removed. The primary research question is based on understanding what needs to happen to create a significant change in rank, not how a change would occur, so a prediction is given for the group of subfactors as a whole without consideration for the correlation between subfactors.

The results of this simulation showed that a school ranked in the mid-30 s could move into the top 20, with the starting point levels shown in Table 4, but this would likely require a corresponding change in its undergraduate academic reputation score. These changes are *not independent* of each other and would all need to happen concurrently either through deliberate action or coinciding changes coming from the correlation between subfactors. For example, in addition to an undergraduate academic reputation score typical of a top-20 school, this university would need to have submitted values in 2012 showing: an acceptance rate close to 25 % or lower, students with SAT scores in the 95th percentile or above, a student to faculty ratio of around 8, a percentage of alumni giving in the high twenties, and a graduation rate of close to 90 %, to highlight a few subfactors. Furthermore, the university would most likely be spending about \$86,000 per student and have an average faculty compensation package (adjusted for cost of living) of about \$150,000.

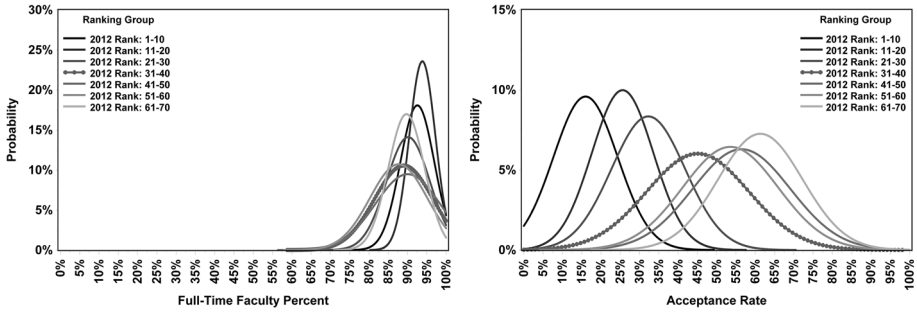


Fig. 7 Exemplar probability-density curves for full-time faculty percent and acceptance rate illustrate how both the weight assigned to a subfactor and the inherent variability within a subfactor, which are both taken into account by the simulations, will impact the amount of change required to move up in rank

Table 4 Subfactor starting levels required to move into the top 20 for a university ranked in the mid-30 s

Criteria	US News 2012 submitted value	Subfactor starting point in order to be ranked in the Top 20	Movement required to move into the Top 20
Peer Assessment	(rated 3.4)	4.2	0.8
High School Counselor Assessment	(rated 3.9)	4.5	0.6
Acceptance Rate	38 %	25 %	−13 %
High School Standing in Top 10 %	79 %	85 %	6 %
Median SAT percentile (score)	90 % (1,300)	95 % (~ 1,380)	5 % (+80)
Median ACT percentile (score)	96 % (~ 30)	97 % (~ 31)	1 % (+1)
Faculty Compensation	\$140,793	\$150,793	\$10,000
Percent faculty with top terminal degree	94 %	96 %	2 %
Percent faculty that is full-time	86 %	94 %	8 %
Student/faculty ratio	10	8	−2
Class size under 20	69 %	70 %	1 %
Class size over 50	11 %	9 %	−2 %
Average graduation rate	84 %	90 %	6 %
Average retention rate	96 %	96 %	0
Financial resources per student	\$73,923	\$85,923	\$12,000
Average alumni giving rate	17 %	27 %	10 %

When accounting for the number of students and faculty at this university, improving these two subfactors alone would require a sustained increase of over \$112,000,000 per year to be allocated. Other required changes, such as decreasing class sizes, increasing graduation rates, or attracting a greater number of highly qualified students (as defined by SAT scores and class standing), would add to the expense of pursuing a ranking change. Finally, the peer assessment score would need to increase 0.8 points, which represents a move that is more than six standard deviations away from the starting point. Based on the small amount of historical change in this subfactor, there is a less than 0.01 % probability that this change would occur. The totality of these changes and the anticipated expenses are very

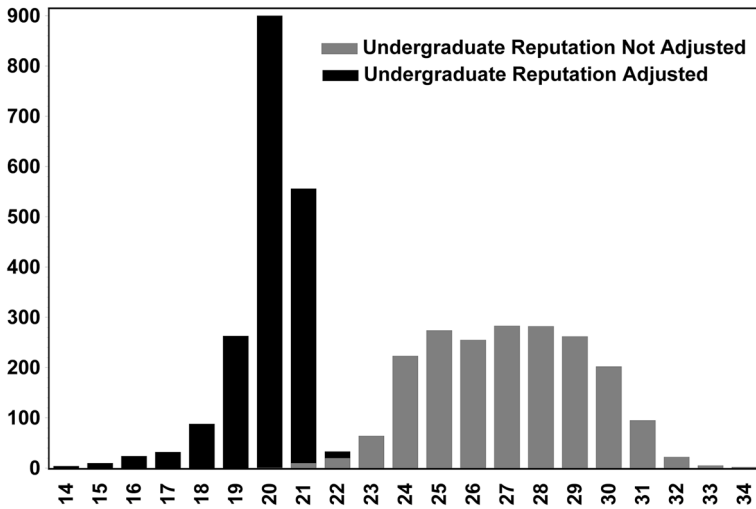


Fig. 8 Examining the distribution of the predicted rankings clearly shows that a change in undergraduate reputation score would be required to move into the top 20

substantial for any university ranked at this level and point to how challenging it would be for a university to move up 15 points into the top 20.

The actual ranks resulting from the simulation of these conditions ranged from 24 to 14, with an overall ranking of 20 occurring most frequently (Fig. 8). Again, because the subfactors were perturbed, the simulations predict a range of ranks. If all of these changes were made, but a corresponding change in undergraduate reputation did not follow, the second simulation showed that a rank between 25 and 30 would be more typical, and this university would never move into the top 20.

Conclusion and Implications

By providing a method to faithfully recreate the *U.S. News* ranking process and establishing a threshold for “noise,” this research establishes the impacts that planned or incidental changes have on a ranking process where the inputs are always varying. As well, it provides empirical evidence of how much change will be required when comparative institutions are not held constant. These results are limited to universities ranked within the *U.S. News* national university category.

The findings have important practical implications for university leaders when setting tactical goals and communicating expectations, and this is particularly true for communications and marketing offices, alumni directors, and university presidents. As shown, when institutions are not purposefully changing the underlying *U.S. News* inputs but instead waver around established levels, changes in ranking should be expected of up to ± 4 points depending on a university’s starting point. With this more precise understanding of the expected “noise” in the *U.S. News* rankings, institutional researchers and university administrators can better communicate the amount of change that could reasonably be anticipated. This may dissipate the pressure to constantly make small adjustments or reevaluate submitted *U.S. News* data, and it can combat the tendency to

manipulate numbers. Even if such practices persist, it should be clear that small changes will have very little impact in the context of the inherent variation in the subfactors themselves.

These results also allude to more general policy considerations. While an improvement in rank may bring positive benefits to a university, this research shows that meaningful rank changes for top universities are difficult and would occur only after long-range and extraordinarily expensive changes, not through small adjustments. Universities might be best served by focusing their efforts and resources on what they do best, not what is being measured by *U.S. News*, by making improvements that are in line with their own goals and are centered around the success of their unique student population.

Three areas could be considered for future research to build on these results. If desired, this same process could be applied to other *U.S. News* ranking categories, such as the national liberal arts university category, to determine if similar conclusions would apply. Next, the accuracy of this model could be improved by determining how to accurately recreate the predicted graduation rate used by *U. S. News*. With this information, a new predicted graduation rate (and graduation rate performance value) could be recalculated as the subfactors were changed during simulations. Finally, this research could be expanded by using this model to predict the impact of subfactors for other lower ranked universities in the top tier that want to improve their rank, yet not as far as the top 20.

This model should not be interpreted as a how-to guide for improving a university's rank but as a tool for making data-driven decisions in order to have meaningful conversations about the *U.S. News* ranking process. The model, and the simulation capability it provides, is useful for communicating the impacts of changes on a university's overall rank. It specifically identifies what would be required for an institution to make a significant ranking change and the influence of any one subfactor alone. It is clearly seen from the example shown that it would take a substantial amount of additional resources, directed in a very purposeful way, to become a top-ranked national university. With these empirically derived requirements, institutional leaders can now weigh the perceived benefit of moving up in rank against the predicted additional costs and the degree to which making changes to pursue a change in rank might alter the very nature of the university.

Acknowledgments The authors would like to thank our manuscript reviewers for their guidance and helpful feedback, especially Andrew Wall and Margaret Kearney.

References

- Adams, W. (2012). A predictable scandal. *Inside Higher Education*. Accessed Oct 12, 2012, from <http://www.insidehighered.com/views/2012/02/07/essay-rankings-scandal-claremont-mckenna>.
- Bastedo, M., & Bowman, N. (2010). U.S. News & World Report college rankings: Modeling institutional effects on organizational reputation. *American Journal of Education*, 116(2), 163–183.
- Betsinger, A. (2009). One ranking to rule them all: modeling U.S. News & World Report's predicted graduation rate, 2008 Research Grant Final Project Report. Accessed Aug 12, 2012, from http://www3.airweb.org/images/final_research_paper_betsinger_2008.pdf.
- Bowman, N., & Bastedo, M. (2009). Getting on the front page: Organizational reputation, status signals, and the impact of U.S. News and World Report on student decisions. *Research in Higher Education*, doi:10.1007/s11162-009-9129-8.
- Clarke, M. (2002). News or noise? An analysis of U.S. News and World Report's ranking scores. *Educational Measurement: Issues and Practice*, 21(4), 39–48.
- Clarke, M. (2004). Weighing things up: A closer look at U.S. News & World Report's ranking formulas. *College and University*, 79(3), 3–9.

- Dichev, I. (2001). News or noise? Estimating the noise in the U.S. News university rankings. *Research in Higher Education*, 42(3), 237–266.
- Ehrenberg, R. (2003a). Reaching for the brass ring: The U.S. News & World Report rankings and competition. *Review of Higher Education*, 26(2), 145.
- Ehrenberg, R. (2003b). Method or Madness? Inside USNWR College Rankings. *Cornell Higher Education Research Institute (CHERI)*. Paper 39. Accessed Sept 26, 2012, from <http://digitalcommons.ilr.cornell.edu/workingpapers/42>.
- Farrell, E., & Van Der Werf, M. (2007). *The Chronicle of Higher Education*. Accessed Sept 26, 2012, from <http://www.chronicle.com/article/Playing-the-Rankings-Game/4451>.
- Grewel, R., Dearden, J., & Llilien, G. (2008). The university rankings game: Modeling the competition among universities for ranking. *The American Statistician*, 62(3), 232–237.
- Griffith, A., & Rask, K. (2007). The influence of the U.S. News and World Report collegiate rankings on the matriculation decision of high-ability students: 1995–2004. *Economics of Education Review*, 26, 244–255.
- Jaschik, S. (2007). Should U.S. News Make Presidents Rich? *Inside Higher Education*. Accessed Sept 26, 2012, from <http://insidehighered.com/news/2007/03/19/usnews>.
- Lederman, D. (2009). ‘Manipulating,’ Er, Influencing ‘U.S. News.’ *Inside higher education*. Accessed June 5, 2009, from <http://www.insidehighered.com/news/2009/06/03/rankings>.
- Machung, A. (1998). Playing the rankings game. *Change*, 30(4), 12–16.
- McDonough, P., Antonio, A., Walpole, M., & Perez, L. (1998). College rankings: Democratized college knowledge for whom? *Research in Higher Education*, 39(5), 513–517.
- Meredith, M. (2004). Why do universities compete in the ratings game? An empirical analysis of the effects of the U.S. News and World Report college rankings. *Research in Higher Education*, 45(5), 443–461.
- Monks, J., & Ehrenberg, R. (1999). The impact of U.S. News & World Report college rankings on admission outcomes and pricing policies at selective private institutions. *Cornell higher education research institute (CHERI)*. Paper 1. <http://digitalcommons.ilr.cornell.edu/cheri/1>.
- Morse, R., & Flanigan S. (2011). How we calculate the rankings. Accessed Jan 17, 2012, from <http://www.usnews.com/education/best-colleges/articles/2011/09/12/how-us-news-calculates-the-college-rankings-2012>.
- Morse, R., Flanigan, S., & Tolis, D. (2011). Association for Institutional Research Forum. May 25. U.S. News & World Report’s America’s Best Colleges: What Happened in the 2011 Edition and What’s Ahead for the New Ranking?
- National Association for College Admission Counseling (2011). Report of the NACAC Ad Hoc Committee on U.S. News & World Report Rankings. Accessed July 11, 2012, from <http://www.nacacnet.org/about/Governance/Comm/Pages/NACACUSNewsAdHocCommittee.aspx>.
- Porter, S. (2000). The robustness of the graduation rate performance indicator used in the U.S. News & World Report college rankings. *The Case International Journal of Educational Advancement*, 1(2), 145–164.
- Sauder, M., & Lancaster, R. (2006). Do rankings matter? The effects of U.S. News & World Report Rankings on the admissions process of law schools. *Law & Society Review*, 40(1), 105–134.
- Schmidt, P. (2008). Most colleges chase prestige on a treadmill, researchers find. *The chronicles of higher education*. Accessed Nov 10, 2012, from <http://chronicle.com/article/Colleges-Chase-of-Prestige/1315>.
- Sponsler, B. (2009). The role and relevance of rankings in higher education policymaking. *Institute for Higher Education Policy Issue Brief, IB06200909*, 1–24.
- Webster, D. (1992). Rankings of undergraduate education in U.S. News & World Report and money: are they any good? *Change*, 24(2), 18–31.
- Webster, T. (2001). A principal component analysis of the U.S. News & World Report tier rankings of colleges and universities. *Economics of Education Review*, 20, 235–244.
- Yablon, M. (2001). Test flight: The scam behind SAT bashing. *New Republic.*, 30, 24–25.