Jacob Fowles
H. George Frederickson
University of Kansas
Jonathan G. S. Koppell
Arizona State University

University Rankings: Evidence and a Conceptual Framework

Jacob Fowles is associate professor in the School of Public Affairs and Administration at the University of Kansas. His research has appeared or is in press in such journals as the *Journal of Policy Analysis and Management, American Review of Public Administration, Administration & Society,* and *Journal of Higher Education*. He received his PhD from the Martin School of Public Policy and Administration at the University of Kentucky.
E-mail: jacob.fowles@ku.edu

H. George Frederickson is Emeritus Distinguished Professor of Public Affairs at the University of Kansas and President Emeritus of Eastern Washington University. He is former president of the American Society for Public Administration and a fellow of the National Academy of Public Administration. He is founding editor of *Journal of Public Administration Research and Theory* and coauthor of *The Public Administration Theory Primer.*
E-mail: gfred@ku.edu

Jonathan G. S. Koppell is dean of the College of Public Service and Community Solutions and Lattie and Elva Coor Presidential Professor at Arizona State University. He is author of *World Rule: Accountability, Legitimacy and the Dynamics of Global Governance* and *The Politics of Quasi-Government.* He is a fellow of the National Academy of Public Administration and previously directed the Millstein Center for Corporate Governance and Performance at Yale University.
E-mail: koppell@asu.edu

Public Administration Review, Vol. 76, Iss. 5, pp. 790–803. © 2016 by The American Society for Public Administration.
DOI: 10.1111/puar.12610.

Abstract: *University ranking has high public visibility, the ranking business has flourished, and institutions of higher education have not been able to ignore it. This study of university ranking presents general considerations of ranking and institutional responses to it, particularly considering reactions to ranking, ranking as a self-fulfilling prophecy, and ranking as a means of transforming qualities into quantities. The authors present a conceptual framework of university ranking based on three propositions and carry out a descriptive statistical analysis of U.S. and international ranking data to evaluate those propositions. The first proposition of university ranking is that ranking systems are demarcated by a high degree of stability, equilibrium, and path dependence. The second proposition links ranking to institutional identity. The third proposition posits that rankings function as a catalyst for institutional isomorphism. The conclusion reviews some important new developments in university ranking.*

Practitioner Points

- Because rankings simplify, decontextualize, and magnify small differences, they can incentivize managers to focus on relative positioning rather than improvement in absolute terms.
- Rankings, once established, become important components of identity, and those organizations privileged by rankings strategically leverage them to shape their institutional identity.
- Ranking systems can function to inhibit diversity and promote uniformity and standardization.

A generation has passed since *U.S. News & World Report* (*USNWR*) published its first newsstand guidebook *America's Best Colleges* in 1983 and *Best Graduate and Professional Schools* in 1987. More than a decade has passed since the Shanghai Jiao Tong University Institute of Education first published the Academic Ranking of World Universities (ARWU) in 2003. In the ensuing years, the college and university ranking business has flourished. "There are three leading global ranking systems plus eight other global rankers of varying significance, and there are over 50 national (U.S, Korea, Germany, Canada, etc.) ranking systems" (Hazelkorn 2011, 5). University ranking systems have proven popular because they enable and formalize comparison. At the university level, it has been determined, for example, that a one-rank improvement in the *USNWR* best colleges ranking leads to a 1 percentage point increase in the number of applicants the following year (Luca and Smith 2013). Because higher education is an important driver of economic development, university rankings also influence state and national competiveness (Hazelkorn 2011).

As the university ranking business has flourished, so, too, has the study of it. Research by university faculty in the formative years of university rankings tended to critique the endeavor (Diamond and Graham 2000; Dichev 2001; Frederickson 2001; Graham and Diamond 1997). In recent years, the study of ranking has become more longitudinally comparative, more analytically sophisticated, and more global (Erkkilä 2013; Espeland and Sauder 2007; Frederickson and Stazyk 2010; Hazelkorn 2011; Jones 2013; Sauder and Espeland 2009). With the passage of time, it is clear that ranking universities is not only here to stay, it is proving to be both resilient and surprisingly influential in university policy and practice.

The ranking of universities is not a distinct phenomenon applicable only within the domain of higher education. Rather, the evolving salience of rankings in higher education is best viewed within a broad context of modern demands for "accountability, transparency, and efficiency … and social measures designed to evaluate the performance of individuals and organizations" (Espeland and Sauder 2007, 1). Indeed, rankings, ratings, and report cards are established and prominent components of initiatives to improve the efficiency and effectiveness of health care providers (Dranove et al. 2003; Fung et al. 2008; Mukamel, Haeder, and Weimer 2014), primary and

secondary schools (Hanushek and Raymond 2004; Kane and Staiger 2002), public sector organizations (Sanger 2008; Van Thiel and Leeuw 2002), and not-for-profit entities (Eckerd and Moulton 2011; Sloan 2009), among others. A rich and developing literature in public administration surveys the practical and theoretical implications of the public's demand for performance measurement. These studies highlight the important and necessarily broad role played by external actors in translating organizational data into relevant and digestible performance information (Hood 2012; Lavertu 2015) and emphasize the potential ramifications of the large degree of latitude and discretion accompanying such efforts (Cooley 2015; Hood, Dixon, and Beeston 2008). If nothing else, the popularity of rankings noted in these diverse literatures reinforces the growing public appetite for such efforts and therefore demonstrates that rankings can only be ignored by organizations at their peril (Van Dyke 2005).

In this article, we describe the rankings of universities as an example of the social construction of reality and of institutionalization (Berger and Luckmann 1966; March and Olsen 1989). University rankings impose on the richly varied, often uneven, and idiosyncratic world of higher education positivist assumptions of order based on formalized hierarchies of quality Ranking applies rationalist assumptions to the often ambiguous and hazy processes and objectives of universities and to the distinctive characteristics of institutions of higher education—public or private, flagship or regional, research or teaching, selective or open, large or small. Once institutionalized, ranking regimes "shape the definitions of alternatives and influence the perception and the construction of the reality within which action takes place" (March and Olsen 1995, 19). Action, in the institutionalized context, is based on established structures, roles, identities, rules, and practices and the logic of appropriateness or "making sense" of the situation. We first review three important features of university ranking from the institutionalist perspective: ranking and reactivity, ranking as self-fulfilling prophecy, and ranking as commensuration (Espeland and Sauder 2007). We then set out a conceptual framework of rankings in higher education designed to facilitate scholarly conversations and debates, research, and further theoretical development.

University rankings impose on the richly varied, often uneven, and idiosyncratic world of higher education positivist assumptions of order based on formalized hierarchies of quality.

Ranking and Reactivity

Reactivity is understood as describing and explaining how "individuals (and organizations) alter their behavior in reaction to being evaluated, observed, or measured" (Espeland and Sauder 2007, 6). Ranking and other forms of performance measurement and organizational evaluation elicit responses from the individuals and organizations being ranked or measured. In many cases, reactivity is keenly understood by those promulgating rankings as an instrument through which they can attempt to alter behavior in the direction of policy they favor. For example, New York State publicly ranks or "scorecards" the performance of medical doctors using criteria such as mortality rates. In response, 79 percent of cardiologists reported that their decision to "perform angioplasty or to intervene in critically ill patients was influenced by potential effects on their scorecards" (Narins et al. 2005, referenced in Espeland and Sauder 2007, 2). Strathern (1997, 2002) found that the application of performance and evaluative measures in Great Britain was central to massive reforms in schooling, universities, budgeting, and health care in the 1980s. In short, ranking and rating has become a primary tool in the advocacy group toolkit (Gormley and Weimer 1999).

In its college ranking program, *USNWR* aims to provide assessable information to each potential education consumer and to assist each in the search for the best school. *USNWR* makes no statements about transformation in higher education or about holding universities accountable. Nevertheless, the initial and continuing reactions to *USNWR* college rankings have set in motion important institutional changes. It is difficult to use the language of institutional improvement or advancement without reference to measures of performance and comparison, particularly comparison by rank. In response to formalized ranking regimes, both the substance and the language of quality in higher education have been reduced to simplified surrogates of quality: rank, status, and prestige.

One need not look hard to identify institutional adaptations to rankings, although whether such efforts are consistent with the spirit of rankings remains subject to some debate. For instance, George Washington University recently abolished the requirement for students to submit an SAT or ACT score for consideration for undergraduate admission. On one hand, this action can be viewed as firing a shot across the bow of the trend of increased quantification within the *USNWR* rankings; on the other, the fact that the number of applicants will undoubtedly rise relative to acceptances after the test score requirement is dropped means that George Washington will likely become a more selective institution in the eyes of *USNWR* after the policy change.

Ranking as Self-Fulfilling Prophecy

Among the more important characteristics of reaction to university ranking is Robert Merton's concept of the self-fulfilling prophecy, "a false definition of the situation evoking a new behavior which makes the originally false definition of the situation come true" (1968, 182–83, 477). The use of the self-fulfilling prophecy concept need not be limited to false beliefs or understandings. Any expectation or assumption that is defined or understood to be real, whether true or false, when subjected to measurement or evaluation, will increase the validity of the original expectation or assumption and thereby encourage behaviors that conforms to it (Espeland and Sauder 2007). For example, *USNWR*'s current ranking metrics assign 12.5 percent of a college's grade or rank to be a function of student selectivity, with 65 percent of that based on student standardized test scores, 25 percent on the proportion of matriculating freshmen in the top decile of their graduating class, and 10 percent on the college's acceptance rate. Empirical evidence suggests that ACT, SAT, and other test scores are weaker predictors of college success than high school grades, and students who do not submit test scores (an increasing number of colleges do not require test scores of applicants) appear to do as well in college as those who do not. Nevertheless, test scores are weighed more heavily than grades in *USNWR* assumptions about student selectivity (Zwick 2007).

Included in student selectivity criteria is the assumption that the fewer students an elite university admits from its applicant pool, the better that university must be.

These student selectivity criteria and the assumptions on which they are based are, at a minimum, debatable. Nevertheless, what is not debatable is that university-level responses to *USNWR* student selectivity criteria have become self-fulfilling prophecies. In pursuit of improved rankings, college admissions processes (and admissions to law schools, graduate schools, etc.) work to generate large pools of applicants, favor applicants with high test scores, and try to limit those admitted to those in the top 10 percent of their high school graduating class (Lovett 2005). Colleges, particularly in the upper ranks, conform to the *USNWR* assumptions about student selectivity and have made them come true. Similarly, the 22.5 percent of the *USNWR* rank determined by student retention provides a compelling explanation for the current trends of hiring of student retention specialists, the opening of offices of student success, and the development of systems for targeting services to students identified as high risk of dropout prior to degree completion.

Those who study the *USNWR* ranking of law schools have found that past rankings are the strongest predictor of current reputation score, consistent with the well-known "anchoring effect" phenomenon uncovered in recent work in psychology and behavioral economics (Kahneman 2011; Tversky and Kahneman 1974). Ample empirical literature demonstrates the powerful conditioning effect of past rankings on current assessments of quality (Bastedo and Bowman 2010; Bowman and Bastedo 2011; Stake 2006). For example, fully 40 percent of law school rankings are based on so-called peer assessment, evaluation of each law school (on a five-point scale from "marginal" to "outstanding") through surveys filled out by deans, faculty members, lawyers, and judges. Because it is impossible for a dean or a judge to know very much about each of the approximately 200 accredited law schools, respondents tend to rely on past judgments of law schools, codified in previous year rankings (Espeland and Sauder 2007).

Ranking as Commensuration

Commensuration is the transformation of qualities into quantities that share a metric. Commensuration is fundamental to quantitative measuring and numeric comparison (Espeland 1998; Espeland and Stevens 1998). "Commensuration shapes what we pay attention to, which things are connected to other things, and how we express sameness and difference" (Espeland and Sauder 2007, 16). The processes of commensuration involve sorting out those characteristics and qualities of universities that are to be included in a shared metric and those to be excluded. It is a process of simplification and decontextualization that reduces the qualities and characteristics of universities to a few and, ultimately, to one number. Ranking is, therefore, an exercise in simplification.

The processes of commensuration involve sorting out those characteristics and qualities of universities that are to be included in a shared metric and those to be excluded.

The processes of simplifications of this type, following March and Simon (1958), make such numeric information seem more authoritative, more robust, and more definitive than narrative information. Simplification masks complexity, obscures assumptions, and hides ambiguity. Information thus simplified is made more portable and more easily recallable. It is much easier to remember a university's rank than to recall the details of a narrative description. It is also easy to assume that the meaning of what appears to be an authoritative number or rank is universal and stable (Espeland and Sauder 2007).

University rankings are constructed using raw scores, such as median grade-point average or median SAT. Such raw scores are obviously quite highly correlated, and the transformation of such continuous measures into ordinal scales magnifies the minute differences between the universities ranked first, second, and third. In *USNWR* rankings, ties and multiple ties are common. Rankings produce a hierarchical relationship between each university being ranked, with seemingly equally sized intervals between universities that are "better than" or "worse than" other universities. In other words, rankings assign precise numbers to each institution although distinctions between institutions are often minuscule (Hood, Dixon, and Beeston 2008). This exercise serves primarily to incentivize universities to focus on relative positioning rather than absolute changes (Espeland and Sauder 2007).

A Conceptual Framework of University Ranking

Having reviewed some patterns of university responses to ranking regimes, the processes of self-fulfilling prophecy, and the effects of commensuration, we employ these axioms to propose a conceptual framework of university ranking. We utilize this framework for the analysis of data and for further descriptions of responses to rankings and to our claim that university ranking regimes constitute an "ivory cage," a higher education version of Max Weber's (1968) "iron cage," a state of institutional isomorphism in which universities are systematically incentivized toward homogeneity (DiMaggio and Powell 1983).

> **Proposition 1:** Under ranking regimes, universities and colleges may in the short run move incrementally up or down the ranks, but in the long run university ranking tends toward equilibrium and system stability.

Among the earliest clues that ranking regimes tend toward stability are found in the work of Dichev (2001). The author analyzed the top 25 national universities and the top 25 national liberal arts colleges for the early years (1989–98) of the *USNWR* rankings, finding that

> [C]hanges in the USN rankings have a strong tendency to revert in the next two rankings. The reversibility in rankings is strong not only in statistical terms but seems to account for a strikingly large part of the total variation in ranking changes. Using a simple model of two-period reversibility, it appears that between 70 and 80 percent of the variation in ranking change is due to noise: transitory effects which quickly disappear in later rankings. Thus, much of the "news" in *USNWR* annual college rankings is essentially meaningless noise. (Dichev 2001, 239)

Here, we update that work and expand on its scope. Following this early analysis of university rankings, table 1 sets out a 12-year compilation of *USNWR*'s ranking of the 50 best U.S. universities starting in 2000, arranged by their rank in 2012, the right-hand column.

Simply scanning table 1 suggests a general equilibrium. Harvard, Princeton, and Yale are ranked either first, second, third, or fourth (Princeton once and Yale once) depending on the year. Columbia is ranked fourth in 2011 and 2012 but was ranked between eighth and eleventh in earlier years. Chicago is nearly the same. Perusing the reported individual components that make up the rankings (not reported here) suggests that these improvements seem to be largely driven by changes in admissions selectivity.

An application of simple longitudinal descriptive statistics is presented in table 2, which displays historical correlation coefficients for two components of the *USNWR* rankings: rank

Table 1 *USNWR* Top 50 University Rankings, 2000–2012 (Sorted by Rank in 2012)

University	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Harvard University	2	2	2	2	1	1	1	2	2	1	1	1	1
Princeton University	4	1	1	1	1	1	1	1	1	2	1	2	1
Yale University	4	2	2	2	3	3	3	3	3	3	3	3	3
Columbia University	10	10	9	10	11	9	9	9	9	8	8	4	4
University of Chicago	13	10	9	12	13	14	15	9	9	8	8	9	5
Stanford University	6	6	5	4	5	5	5	4	4	4	4	5	5
Massachusetts Institute of Technology	3	5	5	4	4	5	7	4	7	4	4	7	5
California Institute of Technology	1	4	4	4	5	8	7	4	5	6	4	7	5
University of Pennsylvania	7	6	5	4	5	4	4	7	5	6	4	5	5
Duke University	7	8	8	4	5	5	5	8	8	8	10	9	10
Dartmouth College	11	9	9	9	9	9	9	9	11	11	10	9	11
Northwestern University	14	13	12	10	11	11	12	14	14	12	12	12	12
Johns Hopkins University	7	15	16	15	14	14	13	16	14	15	14	13	13
Washington University (St. Louis)	17	15	14	12	9	11	11	12	12	12	12	13	14
Brown University	14	15	16	17	17	13	15	15	14	16	16	15	15
Cornell University	11	10	14	14	14	14	13	12	12	14	15	15	15
Vanderbilt University	20	22	21	21	19	18	18	18	19	18	17	17	17
Rice University	14	13	12	15	16	17	17	17	17	17	17	17	17
University of Notre Dame	19	19	19	18	19	18	18	20	19	18	20	19	19
Emory University	18	18	18	18	18	20	20	18	17	18	17	20	20
University of California, Berkeley	20	20	20	20	21	21	20	21	21	21	21	22	21
Georgetown University	23	23	22	24	23	25	23	23	23	23	23	21	22
Carnegie Mellon University	23	23	22	21	23	22	22	21	22	22	22	23	23
University of Southern California	42	35	34	31	30	30	30	27	27	27	26	23	23
Wake Forest University	28	28	26	25	28	27	27	30	30	28	28	25	25
University of Virginia	22	20	24	23	21	22	23	24	23	23	24	25	25
University of California, Los Angeles	25	25	26	25	26	25	25	26	25	25	24	25	25
University of Michigan–Ann Arbor	25	25	25	25	25	22	25	24	25	26	27	29	28
Tufts University	29	29	28	28	27	28	27	27	28	28	28	28	29
University of North Carolina at Chapel Hill	27	25	28	28	29	29	27	27	28	30	28	30	29
Brandeis University	31	31	34	31	32	32	34	31	31	31	31	34	31
Boston College	39	38	38	40	40	37	40	34	35	34	34	31	31
College of William and Mary	29	30	30	30	31	31	31	31	33	32	33	31	33
New York University	34	33	32	35	35	32	37	34	34	33	32	33	33
University of Rochester	32	33	36	36	35	37	34	34	35	35	35	37	35
Georgia Institute of Technology	40	35	41	38	37	41	37	38	35	35	35	35	36
University of California, San Diego	32	31	31	31	32	35	32	38	38	35	35	35	37
University of Miami (FL)	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	50	47	38
University of California, Davis	42	41	41	43	43	42	48	47	42	44	42	39	38
Lehigh University	34	38	38	40	37	37	32	33	31	35	35	37	38
Case Western Reserve University	34	38	38	37	37	35	37	38	41	41	41	41	38
University of Washington	44	45	45	47	45	46	45	42	42	41	42	41	42
University of Wisconsin–Madison	34	35	32	31	32	32	34	34	38	35	39	45	42
University of California, Santa Barbara	44	45	48	47	45	45	45	47	44	44	42	39	42
University of Texas at Austin	44	49	48	47	N/R	46	N/R	47	44	47	47	45	45
Yeshiva University	44	45	41	40	40	46	45	44	N/R	50	N/R	50	45
University of California, Irvine	49	41	41	45	45	43	40	44	44	44	46	41	45
Pennsylvania State University–University Park	40	44	46	45	48	50	48	47	48	47	47	47	45
University of Illinois Urbana-Champaign	34	41	36	38	40	37	42	41	38	40	39	47	45
Rensselaer Polytechnic Institute	N/R	49	48	47	48	46	43	42	44	41	42	41	50
Tulane University	44	45	46	43	44	43	43	44	50	N/R	50	N/R	50
George Washington University	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	50
Syracuse University	N/R	N/R	N/R	N/R	N/R	N/R	50	N/R	50	N/R	N/R	N/R	N/R
University of Florida	49	N/R	N/R	N/R	48	50	50	47	49	49	47	N/R	N/R
Pepperdine University	N/R	49	48	47	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R
Texas A&M University–College Station	N/R	N/R	48	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R

N/R = Not ranked.

Table 2 Longitudinal Correlations of *USNWR* Ranks and Scores

	2012 Component	
	Rank	*Overall Score*
2011	0.994	0.997
2010	0.990	0.995
2009	0.986	0.993
2008	0.983	0.990
2007	0.979	0.987
2006	0.967	0.982
2005	0.971	0.982
2004	0.973	0.979
2003	0.968	0.976

(1–50, inclusive of ties) and overall score (which *USNWR* scales so that the top-ranking university in a given year receives a score of 100). Each row reports the estimated coefficient of determination for the value of the *USNWR* score component in the year listed in column 1 and the corresponding value for the year 2012, calculated using the 47 universities that are observed as Tier 1 institutions in each year observed in the data. We report the Pearson's correlation coefficient for overall score and the Spearman's rank correlation coefficient for ranks.

As this table demonstrates, there are strong statistical associations between the individual components of the *USNWR* rankings over time. As one would expect, the explanatory power of historical values declines slightly as the time lag increases, although a nearly perfect positive linear relationship persists between the 2012 components and the corresponding values from 2003—nine years earlier.

Longitudinal equilibrium as the first proposition of university ranking has also been the primary finding of studies of American law schools (Espeland and Sauder 2007; Jones 2013; Sauder and Espeland 2009; Sauder and Lancaster 2006), schools of business (Iacobucci 2013), and schools of public policy and administration (Frederickson and Stazyk 2010). These findings are particularly relevant to the study of university rankings and ranking systems given the observed "halo effect," in which the reputational assessments of individual programs or departments are conflated by the reputation of the university of a whole, leading, in some cases, to nonexistent professional schools at prestigious universities receiving high reputational assessments (Brooks 2005; Webster 1981).

We turn now to the analysis of global ranking systems. We have chosen the Academic Ranking of World Universities rankings because they have the longest history of global ranking, having started in 2003. The ARWU rankings are useful for comparative purposes because they emphasize research productivity and do not include subjective peer review or student selectivity criteria. ARWU ranking criteria assign 10 percent to alumni winning Nobel Prizes and Fields Medals and 20 percent to staff winning the same awards, 20 percent to highly cited (the "citation impact factor") research in 21 broad academic subjects, 20 percent for papers published in *Nature* and *Science* or other leading peer-reviewed journals, 20 percent to articles in journals indexed in the Science or Social Science Citation Indices, and 10 percent to per capita academic performance (the weighted scores of the foregoing five indicators divided by the number of full-time equivalent [FTE] academic staff). Table 3 presents the ARWU rankings from 2003 to 2014.

The most notable similarity between the *USNWR* ranking of American universities and the ARWU ranking of global universities is the ultra-stability of the top universities in both rankings. It is notable that top American universities are reordered not just by the inclusion of non-American universities but also by the research emphasis found in the ARWU criteria. For example, Duke is number 10 in the *USNWR* rankings and number 36 in the ARWU rankings. Dartmouth is number 11 in the *USNWR* rankings but is not in the top 50 in the ARWU rankings. In the opposite direction, the University of California, San Diego is ranked 37 in the *USNWR* rankings and 14 in the ARWU rankings. The point is that the internal validity of rankings is directly related to the choice of ranking criteria and the assumptions on which these criteria are based. Once ranking criteria are established, however, stability and equilibrium take over.

Table 4 presents an application of the same descriptive statics to the ARWU top 50 global universities that were applied to the *USNWR* rankings of American universities shown in table 2. To maintain comparability with the table 2, table 4 displays the same years and is limited to the 39 universities that appear in the top 50 in all of the years listed.

As was found in the analysis of *USNWR* rankings, the longitudinal correlations among the ARWU ranking components are quite high: the correlations of the 2003 and 2012 components range from .93 to .96.

In the longitudinal analysis of both American and global universities, it is notable that tables 1 and 3 suggest that year-to-year fluctuation in ranking is greater among lower-ranked programs compared with those receiving higher rankings, a proposition that is wholly expected given the literature on anchoring effects in assessment and ranking regimes. We test this observation empirically by analyzing the comparative stability of rankings over time using the *USNWR* rankings of the top 100 (inclusive of ties) American universities for the years 2004–12. Years prior to 2004 are omitted because *USNWR* only assigned ranks and overall scores to the top 50 universities in those years. We begin by calculating year-by-year changes in rankings for each university that appears in the top 100. For instance, MIT was ranked 4 in 2004 and 5 in 2005, so it is assigned a score of –1 in 2005, indicating that it dropped one position. Virginia Tech was ranked 77 in 2007 and 71 in 2008, so it is assigned a score of 6 in 2008, indicating its increase in the rankings. We take the absolute values of these scores, as we are interested in absolute rather than directional changes. Finally, we divide the top 100 universities in each year into quartiles based on the overall score assigned by *USNWR* in that year and calculate the average absolute scores for each quartile across all years. These average scores are shown in table 5.

These scores are interpreted as the expected change (increase or decrease) in rank from for one year to the next for a typical university in each quartile. As the data show, the average member of the top quartile can expect its rank to fluctuate by less than one position each year. As we decrease in ranking to the fourth quartile, the expected annual change increases significantly, with the average member of the bottom quartile experiencing annual changes of over five positions.

Table 3 ARWU Top 50 University Rankings, 2003–14 (Sorted by Rank in 2014)

University	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Harvard University	1	1	1	1	1	1	1	1	1	1	1	1
Stanford University	2	2	3	3	2	2	2	3	2	2	2	2
Massachusetts Institute of Technology	6	5	5	5	5	5	5	4	3	3	4	3
University of California, Berkeley	4	4	4	4	3	3	3	2	4	4	3	4
University of Cambridge	5	3	2	2	4	4	4	5	5	5	5	5
Princeton University	7	7	8	8	8	8	8	7	7	7	7	6
California Institute of Technology	3	6	6	6	6	6	6	6	6	6	6	7
Columbia University	10	9	7	7	7	7	7	8	8	8	8	8
University of Chicago	11	10	9	8	9	9	9	9	9	9	9	9
University of Oxford	9	8	10	10	10	10	10	10	10	10	10	9
Yale University	8	11	11	11	11	11	11	11	11	11	11	11
University of California, Los Angeles	15	16	14	14	13	13	13	13	12	12	12	12
Cornell University	12	12	12	12	12	12	12	12	13	13	13	13
University of California, San Diego	14	13	13	13	14	14	14	14	15	15	14	14
University of Washington	16	20	17	17	16	16	16	16	16	16	16	15
University of Pennsylvania	18	15	15	15	15	15	15	15	14	14	15	16
The Johns Hopkins University	24	22	19	20	19	20	19	18	18	17	17	17
University of California, San Francisco	13	17	18	18	18	18	18	18	17	18	18	18
Swiss Federal Institute of Technology Zurich	25	27	27	27	27	24	23	23	23	23	20	19
University College London	20	25	26	26	25	22	21	21	20	21	21	20
University of Tokyo	19	14	20	19	20	19	20	20	21	20	21	21
University of Michigan–Ann Arbor	21	19	21	21	21	21	22	22	22	22	23	22
Imperial College of Science, Technology and Medicine	17	23	23	23	23	27	26	26	24	24	24	22
University of Toronto	23	24	24	24	23	24	27	27	26	27	28	24
University of Wisconsin–Madison	27	18	16	16	17	17	17	17	19	19	19	24
Kyoto University	30	21	22	22	22	23	24	24	27	26	26	26
New York University	N/R	32	29	29	30	31	32	31	29	27	27	27
Northwestern University	29	30	31	33	29	30	30	29	30	30	30	28
University of Illinois Urbana-Champaign	45	25	25	25	26	26	25	25	25	25	25	28
University of Minnesota, Twin Cities	37	33	32	32	33	28	28	28	28	29	29	30
Duke University	32	31	32	31	32	32	31	35	35	36	31	31
Washington University in St. Louis	22	28	28	28	28	29	29	30	31	31	32	32
Rockefeller University	28	29	30	30	30	32	32	34	33	32	34	33
University of Colorado Boulder	31	34	35	34	34	34	34	32	32	33	33	34
Pierre and Marie Curie University—Paris 6	N/R	41	46	45	39	42	40	39	41	42	37	35
University of North Carolina at Chapel Hill	N/R	N/R	N/R	N/R	N/R	38	39	41	42	41	43	36
University of British Columbia	35	36	37	36	36	35	36	36	37	39	40	37
University of Manchester	N/R	N/R	N/R	50	48	40	41	44	38	40	41	38
University of Texas at Austin	47	40	36	39	38	39	38	38	35	35	36	39
University of Copenhagen	N/R	N/R	N/R	N/R	46	45	43	40	43	44	42	39
University of California, Santa Barbara	26	35	34	35	35	36	35	32	33	34	35	41
University of Paris Sud (Paris 11)	N/R	48	N/R	N/R	N/R	49	43	45	40	37	39	42
University of Maryland, College Park	N/R	N/R	47	37	37	37	37	36	38	38	38	43
University of Melbourne	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	44
University of Texas Southwestern Medical Center at Dallas	34	36	38	38	39	41	48	49	N/R	48	46	45
University of Edinburgh	43	47	47	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	45
Karolinska Institute	39	46	45	48	N/R	N/R	50	42	44	42	44	47
University of California, Irvine	44	N/R	47	44	45	46	46	46	48	45	45	47
Heidelberg University	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	49
University of Munich	48	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	49
Australian National University	49	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R
Utrecht University	40	39	41	40	42	47	N/R	50	48	N/R	N/R	N/R
Rutgers, The State University of New Jersey	38	44	43	46	47	N/R	N/R	N/R	N/R	N/R	N/R	N/R
University of Southern California	40	48	50	47	50	50	46	46	46	46	47	N/R
Technical University Munich	N/R	45	N/R	N/R	N/R	N/R	N/R	N/R	47	N/R	50	N/R
Brown University	49	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R
Vanderbilt University	32	38	39	41	41	42	41	N/R	N/R	50	49	N/R
Pennsylvania State University–University Park	40	43	39	42	43	42	45	43	45	49	N/R	N/R
University of Zurich	45	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R
University of California, Davis	36	42	41	42	43	48	49	46	48	47	47	N/R
University of Pittsburgh	N/R	48	43	48	49	N/R	50	N/R	N/R	N/R	N/R	N/R

N/R = Not ranked.

At first glance, this finding appears to provide empirical evidence that counters the stasis demonstrated earlier. However, the two are easily reconciled if the fluctuation observed in the lower ranks is random noise rather than consistent upward or downward trending over time. In other words, the ranking system overall exhibits a high degree of aggregate stability across all institutions—but the lower-ranked institutions are ordered with less precision in any particular year.

Our findings corroborate those of Gnolek, Falciano, and Kuncl (2014), who developed a dynamic model of the categories (such

Table 4 Longitudinal Correlations of ARWU Ranks and Scores

	2012 Component	
	Rank	*Overall Score*
2011	0.998	0.999
2010	0.986	0.998
2009	0.994	0.997
2008	0.991	0.995
2007	0.987	0.994
2006	0.982	0.993
2005	0.984	0.993
2004	0.975	0.987
2003	0.930	0.961

Table 5 Stability in *USNWR* Rankings, by Quartile

Quartile (Position)	Average Change Score
1 (1–25)	.872
2 (26–50)	1.85
3 (51–75)	3.56
4 (75–100)	5.30

as student selectivity), subfactors (such as SAT or ACT scores), and the weights assigned to them (7.5 percent) by *USNWR* for the years leading up to and including 2012. They determined that annual rank changes four or fewer positions should be considered "noise," with only longer-range changes in the same direction being significant. Using their model, Gnolek, Falciano, and Kuncl demonstrate that for a school ranked in the mid-30s to move into the top 20 would require a sustained increase of more than $112 million to its annual academic expenditures over several years, and even then, because of the small amount of change over time in *USNWR* ranking subfactors, there is less than a .01 percent probability of improved ranking.

> **Proposition 2:** Universities strategically employ the status afforded by rankings to shape and define institutional identity.

As university ranking emerged in the late twentieth century, so, too, did university strategic planning. The parallel trend is not surprising. The language and logic of ranking are suited to the rationalist assumptions of strategic planning. The coming together of ranking and university strategic planning is particularly interesting, especially in the cases of those strategic plans that aspire to significantly improved ranking. Among the more interesting examples of this is found in the case of the University of Kentucky and its former president Dr. Lee T. Todd, Jr.

After a successful career as a professor of electrical engineering and a noted entrepreneur, Dr. Todd took office as the 11th president of the University of Kentucky in 2001. Four years earlier, in 1997, the Kentucky state legislature had made a compact with the University directing it to become a top-20 public research institution by 2020, a classic example of leveraging rankings for the purposes of setting performance targets (Hood 2012). In response, President Todd and his staff developed the "University of Kentucky Top 20 Business Plan" (TOP 20 Plan), which included increased enrollments, increased graduation rates, increased numbers of faculty, increased research funding, and increased university engagement in schools, businesses, farms, and communities. To achieve these goals, the TOP 20 Plan requested an increase of $260 million in state funding over 15 years. At the time, the business plan was adopted, the University of Kentucky was ranked by *U.S. News & World Report* 35th among public research universities and in the unranked second tier in the overall rankings. State appropriations for public higher education increased in absolute terms in Kentucky after the adoption of the compact in 1997, even after adjusting for inflation. However, the trend line becomes relatively flat if these changes are scaled by changes in either total state population or personal income.

After nearly 10 years in office, President Todd resigned in September 2010. At the time of his resignation, the University of Kentucky was ranked 38 by *U.S. News and World Report* among American public universities and 129 overall, suggesting little relative change and a failure to meet the compact's objectives. However, despite the failure to improve meaningfully in rankings, the university did improve in absolute terms in some areas, including peer scores and graduation rates. Under the leadership of the current president, Dr. Eli Capilouto, the new University of Kentucky strategic plan, "Seeing Tomorrow," does not explicitly refer to rankings in framing and describing university goals.

In her analysis of global university ranking systems, Hazelkorn found that "one of the first places where the influence of rankings can be seen is in university vision or mission statement and strategic plans." She notes four common approaches to the incorporation of rankings within strategic planning: "1) rankings as an explicit goal; 2) rankings as an implicit goal; 3) rankings for target setting; and 4) rankings as a measure of success" (2011, 97). The usual form of explicit ranking goals took the form of plans to "be in the top 20" or "be in the first tier." With the passage of time, as the stability of rankings has become increasingly clear, ranking statements as part of university strategic plans tend to describe implicit and more general goals such as "achieving national standing" or being "world-class." Rankings as measures of success now almost always take the form of implicit vision or goal statements, such as "the university is moving in the right direction" or "the university is making progress," rather than describing a specific rank or ranking aspiration.

However, these suppositions are difficult to demonstrate empirically. Little systematic evidence exists that explains the relative prominence and consideration afforded rankings in the strategic planning process. Gathering reliable quantitative data that speak to the explicit utilization of rankings and ranking criteria in guiding university planning and policy making is difficult, as public access to artifacts such as planning documents, minutes from board meetings, and the like is highly uneven across the individual universities. In lieu of broad and comprehensive access to such documents, we instead develop a proxy measure of rankings prominence by taking raw counts of the number of unique web pages that explicitly reference rankings across all pages indexed within a given university's website. We

Little systematic evidence exists that explains the relative prominence and consideration afforded rankings in the strategic planning process.

use this measure to proxy the relative consideration that different universities give to rankings as a means to shape their identities. The quantitative analysis of text scraped from individual websites and the indices maintained by search engines is becoming an increasingly common approach to measuring difficult to quantify social phenomena ranging from the relative priorities of U.S. senators (Grimmer 2010) to citizen racial attitudes (Stephens-Davidowitz 2014) and public perceptions of private firms (Bollen, Mao, and Zeng 2011). We argue that differences in these counts meaningfully demonstrates the validity of proposition 2, in the broader sense that they reflect, with noise, the prominence afforded rank by institutions in defining and shaping their outward persona—a key component of strategic planning.

To gather these data, we begin with the universe of 268 universities listed in the "Best National Universities" list published by *USNWR* in 2012 (the last year of our rankings data), including all four tiers. We first conducted 268 Google web queries to ascertain the root domain for each university's website. For example, the University of Kansas's root domain is "ku.edu," indicating that all pages hosted on all subdomains by the university contain a reference to this common top-level domain name.

We next conducted a second round of 268 Google searches, which were restricted to web pages associated with a single university's root domain. For this round of searches, we limit our results to pages containing the phrases "news and world report rankings" or "news & world report rankings" (thereby avoiding the potential complication associated with "U.S." versus "US"). Each Google search returned an aggregate number of identified search results, indicating the total number of unique web pages hosted on a university's domain that matched the specified search parameters. It is this number that we employ in the analysis that follows. Figure 1 plots the distribution of this variable using a kernel density plot.

As this plot shows, our variable exhibits a high degree of positive skewness, consistent with the expected distribution of a variable measuring raw counts. Accordingly, we report nonparametric

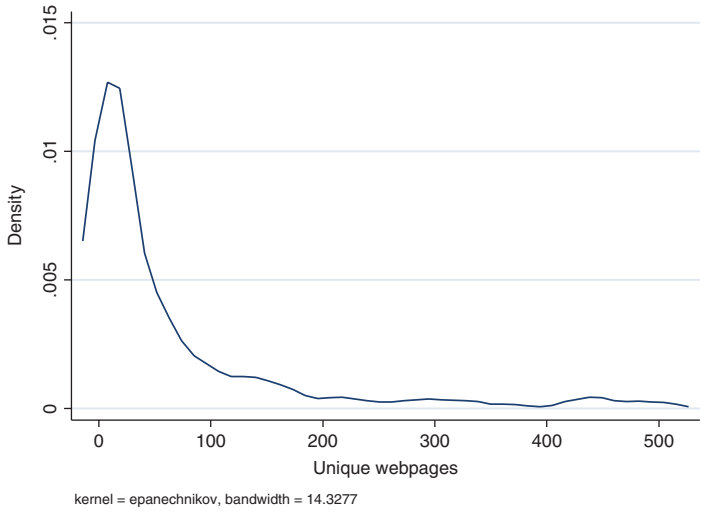

Figure 1 Kernel Density Plot of Unique Web Pages Referencing *USNWR* Rankings

Table 6 Web Pages Referencing *USNWR* Rankings, Disaggregated by *USNWR* Ranking Quintiles

Rank Quartile	Median Number of Web Pages Referencing *USNWR* Rankings	Number of Domains with no Web Pages Referencing *USNWR* Rankings
1	120.5	0
2	28.0	3
3	9.0	10
4	3.5	18

descriptive statistics and employ nonparametric statistical tests in the analysis which follows.

Across all 268 universities, the median number of unique web pages that reference *USNWR* rankings is 25, with an interquartile range of 78.5. Consistent with figure 1, this demonstrates there is marked heterogeneity in terms of the frequency of reference to *USNWR* rankings across university websites. Further, we find systematic differences in the frequency of unique pages making reference to rankings according to the university's *USNWR* rank. Table 6 tabulates the median number of unique web pages referencing *USNWR* rankings and the count of domains that do not reference *USNWR* rankings, disaggregated by *USNWR* rank quartile. Quartile 1 includes universities ranked 1–68. Quartile 2 contains universities ranked 71–132. Quartile 3 includes universities ranked 138–194. Quartile 4 contains all unranked universities (those appearing in Tier 4).

As this table demonstrates, there are significant differences in the median number of web pages containing references to the *USNWR* rankings across the rank quintiles, with *USNWR* rankings being referenced much more frequently on the websites of universities that are of higher *USNWR* rank. A Kruskal-Wallis test of medians strongly rejects the null of no differences in the medians across the specified ranking quartiles ($chi^2(3) = 98.425$, $p = .0001$). Universities in the top quartile are much more likely to reference their ranking than universities of lower rank. This observation is reinforced by the raw counts of domains that do not make mention of *USNWR* rankings. No domains in the top quartile fail to mention rankings, while more than 27 percent of university domains in the lowest quartile do not mention rankings at all. Consistent with proposition 2, this is strong suggestive evidence that highly ranked universities employ the language of rankings and reference their *USNWR* ranks to reinforce their own stature and identity. Patterns of reaction to the work of ranking organizations are the attempts of university leaders to leverage and make sense of the rankings.

Naturally, this evidence is suggestive at best. A plausible alternative to this proposition is that rankings are descriptive rather than proscriptive—that is, the rankings reflect what the privileged universities do rather than independently and proscriptively shape their behaviors. Both are likely true in practice, as the viability of any ranking system in the marketplace hinges on the broader social recognition of the rankings as legitimate and valid (Dill and Soo 2005). There is ample evidence that *USNWR* has been responsive to the market, including most obviously its shift away from a purely reputational survey of quality in its early years to a model that balanced subjective reputational data with more "objective" measures (Meredith 2004). Nonetheless, there is much evidence that suggests that the instrument of the rankings influences university

behaviors in ways largely independent from the underlying quality measures that the rankings intend to measure and evaluate. For example, Ehrenberg (2002) provides an account of a university that proposed to increase faculty salaries for the sole purpose of improving the "faculty resources" metrics of the *USNWR* rankings, independent of any actual discussion of the theoretical or actual improvements in teaching quality that could result. As Bowman and Bastedo lament, "over time, rankings increasingly *become* reputation, rather than reputation being an independent indicator that rankings can use to assess changes in quality" (2011, 432).

> **Proposition 3:** Formalized ranking strengthens the forces of institutional isomorphism. This ultimately incentivizes universities to resemble one another, holding other factors constant.

It was Weber (1968) who described modern organizations as change-resistant "iron cages" on which we come to depend. In their adaption of Weber's iron cage thesis, DiMaggio and Powell (1983) set out a theory of institutional isomorphism that is particularly suited to explaining the influence of university ranking regimes. Modern universities are bureaucracies subject to the forces of stasis, homogenization, equilibrium, and path dependence. Patterns of organizational change associated with responding to university ranking regimes result in "processes that make organizations more similar without making them more efficient" (DiMaggio and Powell 1983, 147).

Theories of institutional isomorphism are particularly suited to explaining the influence of ranking on universities.

Theories of institutional isomorphism are particularly suited to explaining the influence of ranking on universities. Modern universities are highly complex bureaucracies of order, reliability, and predictability. Patterns of institutional isomorphism are a particular characteristic of organizations, such as universities, that operate in fields in which there is an uncertain and difficult to measure relationship between means and end and in fields in which agreement on preferred outcomes is illusive (DiMaggio and Powell 1983). Three forms of institutional isomorphism are evident in higher education. Coercive isomorphism results from imposed rules and policies, a common legal environment, licensing, and especially systems of accreditation. Normative isomorphic pressures include the filtering processes of formal education and credentialing, with professional organizations serving as carriers of norms, standards, and culture. Most important in higher education is mimetic isomorphism in which "organizations tend to model themselves after similar organizations in their field that they perceive to be more legitimate or successful (DiMaggio and Powell 1983, 152). In the case of higher education, one must add the word "prestigious" to the words "legitimate" and "successful" to account for mimicking in higher education.

In his study on the relationship between diversity and reputation in higher education, Van Vught (2008) contrasts conflicting theoretical arguments regarding whether differentiation or homogenization should naturally emerge from systems of higher education. That work focuses on the nature of the dynamics of the relationship between organizations and the environments in which they operate as a key driving factor, identifying the common values held by administrators and faculty members trained and socialized within academia (a form of normative isomorphism) as well as the centralized planning mechanisms adopted in many states and nations (a form of coercive isomorphism) as external environmental factors promoting system stability and driving de-differentiation. To that list, we would add the influence of rankings. Because rankings favor certain metrics that privilege certain activities (academic and research performance) and institutions over others, rankings promote homogeneity by incentivizing universities to focus their efforts and attention in uniform ways to maintain legitimacy and mimic those institutions deemed "high achieving," thereby reinforcing the patterns of "academic drift" (Berdahl 1985; Morphew and Huisman 2002).

To demonstrate the isomorphic properties of rankings, we again utilize the *USNWR* ranking data from 2012, focusing on the 201 unique universities receiving a numeric ranking in that year. To these data, we merge data on the number of baccalaureate degrees issued by each university by Classification of Instructional Programs (CIP) code, a taxonomy developed by the National Center for Education Statistics that aggregates individual fields of study into broader categories based on subject matter. We use the two-digit CIP codes, which recognize 54 unique major fields of study that range from the social sciences, to engineering, to the liberal arts and sciences. In order to avoid issues of scale, we transform each count into a percentage by dividing it by the total number of baccalaureate degrees issued. Table 7 lists all of the 54 CIP codes represented in our data and provides descriptive statistics for each of those degree fields based on our sample of 201 ranked universities.

As one would expect, business, the social sciences, and engineering make up large proportions of the typical university's undergraduate programs. Beyond these three categories, however, there is marked diversity in the composition of *USNWR*-ranked universities' undergraduate programs.

In order to test for a relationship between relative rank and undergraduate program composition, we first calculate pairwise Gower's dissimilarity coefficients for each unique pair of universities. Gower's coefficient is a standard measure for evaluating the similarity of two observations based on a set of specified variables and is widely used in such diverse fields as ecology and computer science (Gower 1971). This measure is bounded by zero, representing two perfectly identical observations, and one, representing two perfectly dissimilar observations.

To determine the relationship between relative rank position and our measure of similarity in undergraduate degree programs, we calculate the absolute difference in rank between the two universities being compared in each unique pair. We then simply regress our dissimilarity coefficients on the absolute rank differences, thereby determining if similarity in undergraduate offerings can be predicted by closeness in *USNWR* rank. We also include in our regression a vector of variables to control for other observed similarities in universities that may have a confounding effect on our estimate of the strength of the relationship between

Table 7 Baccalaureate Degree Share by CIP Code for *USNWR*-Ranked Universities, 2012

	Mean	SD	Min.	Max.
Agriculture, agriculture operations, and related sciences	1.20	2.46	0.00	12.60
Natural resources and conservation	1.15	2.41	0.00	28.34
Architecture and related services	1.04	2.20	0.00	20.27
Area, ethnic, cultural, gender, and group studies	0.81	1.26	0.00	7.39
Communication, journalism, and related programs	5.07	4.28	0.00	19.93
Communications technologies/technicians and support services	0.05	0.23	0.00	2.10
Computer and information sciences and support services	2.26	2.53	0.00	16.19
Personal and culinary services	0.01	0.05	0.00	0.43
Education	3.98	4.79	0.00	28.12
Engineering	10.10	13.10	0.00	87.93
Engineering technologies and engineering-related fields	0.83	1.97	0.00	14.70
Foreign languages, literatures, and linguistics	1.60	1.23	0.00	6.68
Family and consumer sciences/human sciences	1.51	2.50	0.00	10.96
Legal professions and studies	0.15	0.56	0.00	5.68
English language and literature/letters	3.18	1.86	0.00	12.87
Liberal arts and sciences, general studies, and humanities	1.63	2.75	0.00	13.95
Library science	0.00	0.00	0.00	0.00
Biological and biomedical sciences	7.26	4.32	0.00	36.78
Mathematics and statistics	1.54	1.39	0.00	9.91
Military technologies and applied sciences	0.00	0.00	0.00	0.00
Multi/interdisciplinary studies	2.00	2.50	0.00	14.40
Parks, recreation, leisure, and fitness studies	1.71	2.13	0.00	8.70
Philosophy and religious studies	0.86	1.07	0.00	11.11
Theology and religious vocations	0.14	1.05	0.00	14.27
Physical sciences	2.16	2.08	0.00	22.84
Science technologies/technicians	0.01	0.06	0.00	0.82
Psychology	6.21	3.15	0.00	20.04
Homeland security, law enforcement, firefighting, and related protective service	1.13	2.01	0.00	11.66
Public administration and social service professions	1.02	1.46	0.00	7.62
Social sciences	11.17	7.87	0.00	41.88
Construction trades	0.04	0.44	0.00	6.27
Mechanic and repair technologies/technicians	0.01	0.07	0.00	0.70
Precision production	0.00	0.00	0.00	0.00
Transportation and materials moving	0.19	1.07	0.00	11.88
Visual and performing arts	4.67	5.35	0.00	68.27
Health professions and related programs	6.91	7.64	0.00	54.70
Business, management, marketing, and related support services	16.21	9.68	0.00	72.22
History	2.22	1.44	0.00	8.62
Observations		201		

undergraduate offerings and ranks. Specifically, we control for whether the universities represented in each unique pair operate in the same sector (public; private, not-for-profit with religious affiliation; private, not-for-profit with no religious affiliation), are located in the same census region, have identical land grant university status, and are categorized within the same Carnegie Classification category (using the 2000 Carnegie Classification scheme). We also control for the absolute difference in the natural log of total enrollments in order to eliminate the potentially confounding effect of similarity in size. Descriptive statistics for our dependent and independent variables drawing on our pairwise data set are presented in table 8. The results of estimating our regression are reported in table 9. In order to account for potential unmeasured heterogeneity, we apply the Huber-White correction to reported standard errors.

Table 8 Descriptive Statistics, Full Pairwise Data Set

	Mean	SD	Min.	Max.
Gower's dissimilarity coefficient	0.10	0.03	0.03	0.24
Absolute rank difference	66.86	47.00	0.00	193.00
Indicator, = 1 if institutions are in same sector	0.43	0.49	0.00	1.00
Indicator, = 1 if institutions are in same Census region	0.15	0.35	0.00	1.00
Indicator, = 1 if institutions have the same land grant designation	0.62	0.48	0.00	1.00
Indicator, = 1 if institutions are in the same Carnegie Classification group	0.51	0.50	0.00	1.00
Absolute difference in the natural log of total enrollment	0.86	0.68	0.00	4.78
Observations	20,100			

As these results show, we find a strong, positive, and highly statistically significant relationship between the undergraduate degree offerings and rank, indicating that, as absolute difference in rank increases, the structure of undergraduate programs also diverge, all else held constant. This indicates that, controlling for all the other variables included in the model, similarly ranked universities tend to have similar undergraduate degree offerings.

Our control variables also demonstrate that other similarities between universities beyond rank are associated with the similarity of undergraduate offerings as well, in the expected directions: pairs of universities with matching institutional characteristics have similar undergraduate offerings, and undergraduate offerings diverge at universities as discrepancies in size (as measured by the natural log of total enrollments) increase. Although not reported here, we ran additional models including the absolute difference of revenue

Table 9 Regression Results Dependent variable: Gower's dissimilarity coefficient of university baccalaureate degree shares

Absolute rank difference	0.0002*** (0.0000)
Indicator, = 1 if institutions are in same sector (public or private, not for profit)	–0.0032*** (0.0004)
Indicator, = 1 if institutions are in same census region	–0.0022*** (0.0006)
Indicator, = 1 if institutions are both land grant colleges/ universities	–0.0043*** (0.0004)
Indicator, = 1 if institutions are in the same Carnegie classification	–0.0041*** (0.0004)
Absolute difference in the natural log of total enrollment	0.0085*** (0.0003)
Constant	0.0892*** (0.0007)
Observations	20,100

Huber-White standard errors in parentheses.
*p<.10; **p<.05; ***p<.01.

per FTE enrolled student as well as the absolute difference in educational expenditures per FTE enrolled student in order to test for the potential additional confounding influence of differences in wealth and expenditure patterns. Those results yielded substantively similar findings to those reported in table 9.

Admittedly, this is only suggestive evidence of the isomorphic pressure exerted by rankings. Isomorphism is a dynamic process by definition, and our reliance on static, cross-sectional data means that, while we can observe a result consistent with the proposition, we are unable to directly observe the process driving the patterns of outcomes that we identify. However, despite the fact that *USNWR* rankings make no distinctions in the rating criteria according to field or diversity of undergraduate degree offerings, our empirical finding is consistent with a general isomorphic pressure to conform that is reinforced by the *USNWR* ranking regime. This is a finding that warrants further investigation and could be straightforwardly extended utilizing panel data.

If ranking drives processes of change in the direction of institutional homogeneity, are the forces of ranking-driven change making universities less effective, as the theory of institutional isomorphism would predict? A recent comprehensive study of American schools of law suggests that this is the case (Jones 2013). It is suggested elsewhere that if equilibrium is the default condition under university and college ranking regimes, universities and colleges will not risk change except changes in the direction of ranking criteria. Rankings are, therefore, "an enemy of college and university creativity and innovation and a form of institutional isomorphism" (Frederickson and Stazyk 2010, 76). Universities tend to resist change, and contextual factors in addition to ranking, such as systems of accreditation, systems of evaluation and performance measurement, and systems of accountability, magnify those tendencies. Although it is difficult to isolate the influence of ranking from the power of context, this should not be taken to deny the homogenizing influence of rankings.

Conclusions

As we enter the second generation of university ranking, it is appropriate to take stock of what we have learned. Ranking systems exhibit a high degree of macro-level predictability, as that ranking organizations churn their annual rankings as they attempt to hold interest or to make news. We demonstrate here that, once a ranking organization sets its ranking criteria, the order of the ranks of universities is quite stable and seldom exhibit long-term change. Despite the equilibrium associated with university rankings, our analysis provides evidence consistent with the proposition that ranking influences university planning and policy, much of that influence in the direction of criteria set by ranking organizations.

Certainly, the correlations we identify fall short of casual proof of these propositions in practice. Nonetheless, we argue that the quantitative, empirical associations we uncover in the data do provide strong suggestive evidence of the influence of ranking systems on the universities they seek to evaluate that can inform the next generation of rankings research. Future studies should take the next logical step of building on the foundation we establish to differentiate the impact of the multiple alternative and confounding forces that influence university behavior, thereby gauging the relative strengths of these forces. Institutions of higher education vary widely in context and in purpose. Ranking systems must take these differences seriously, as must those doing research on the ranking of institutions of higher education.

There are some important implications of our findings with respect to the continued viability of systems of higher education in the United States and elsewhere that are worthy of discussion. The confluence of evidence suggests that imposition of ranking regimes is not a catalyst of innovation in higher education unless that innovation favors ranking criteria. Rankings simplify, reducing university qualities, contexts, and unique characteristics to simpler metrics. Rankings decontextualize. Rankings magnify small differences. Rankings mask complexity and hide ambiguity. Rankings replace "different than" with "better than." But, even knowing this, university ranking regimes continue to be influential. Why?

The confluence of evidence suggests that imposition of ranking regimes is not a catalyst of innovation in higher education unless that innovation favors ranking criteria.

In the American context, the criteria used by *USNWR* certainly represent the dominant identifying characteristics of the traditional American research university. To be sure, there are variations on that model, including the American regional universities, the urban universities, the liberal arts colleges, the community colleges. These are institutions that, taken together, educate far more people than those educated at the institutions categorized in the Carnegie Classifications as "Doctoral/Research University—Extensive." Nevertheless, research universities continue to define the model privileged by ranking organizations, and thereby rankings continue to advantage historically elite institutions.

Second-generation university ranking is increasingly marked by competing ranking organizations using different criteria. In the United States, this is most notable in the ranking of business schools and MBA programs, with three competing ranking organizations. It is also a characteristic of international or global university ranking regimes, also with three primary ranking organizations. Some

ranking organizations, most notably *Bloomberg Businessweek,* use survey-based methodology and "soft" criteria, such as "student experience," in their rankings of business schools. As with the ranking systems discussed earlier, these approaches demonstrate aggregate stability with micro-level fluctuation that seems attributable to randomness as much as systematic improvement or decline despite the differing methodology employed. For example, the top 15 business schools are always the top 15, but Harvard moves from first to eighth and Duke moved from sixth to first in the 2014 *Businessweek* rankings. Despite the rather obvious question of how or whether the Harvard MBA program got significantly worse and the Duke MBA got significantly better, all in the space of 12 months, this churning, or noise, seems to capture the attention of the media, much like the attention given the ranking of university football teams. More to the point, the MBA ranking program at *Bloomberg Businessweek* is vigorously competing with *USNWR,* the dominant player in the ranking game, and with *Forbes,* the other important MBA ranking game participant.

Some interesting and important recent developments have also emerged that have the potential to introduce significant changes to the rankings landscape. One is the renewed focus on value-added approaches to evaluating institutions of higher education (Cunha and Miller 2014). Value-added measures offer the potential to overcome some of the well-known shortcomings of the traditional ranking systems. First, they largely focus on intermediate or long-term outcomes rather than inputs (such as faculty salaries, volumes of books and periodicals housed, or average ACT score of matriculating freshmen) or outputs (such as grants awarded, articles published, or students retained). Value-added measures instead focus on measures such as degree completion or the economic success of graduates, measured through placements or salaries. Second, value-added measurement attempts to partition these outcomes into two distinct components by explicitly controlling for the relative differences in inputs across universities (in terms of both students as well as other resources), thereby making a more plausible attempt to estimate the direct, marginal contribution of the university to the well-being of the students it serves. Rankings of universities built around value-added measures contribute a very different perspective than that represented by traditional ranking regimes, largely concluding that the highest-achieving institutions are those focusing on engineering, STEM (science, technology, engineering, and math), and medical fields (Rothwell 2015). Proponents of value-added measures argue that they put all universities on equal footing for comparison and that value-added measures of higher education institutions exhibit a high degree of consistency over time. Critics argue that value-added measures simply substitute new biases for old ones, pointing out that differences across universities in tabulating student outcome data combined with relying on noisy outcome measures such as graduate salaries make value-added measures less reliable and useful than they appear.

A related second development in the domain of rankings is the recent attempt by the U.S. federal government to develop an alternative source of institutional rankings. In 2013, President Barack Obama announced that the federal government would develop a new consumer-driven ratings system explicitly designed to compete with *USNWR,* but with a primary focus on college affordability and value (Obama 2013). However, the rankings that the president proposed never materialized, partly because of a lack of support from key institutional stakeholders who questioned both the overall usefulness of the exercise in an already crowded space as well as the validity and reliability of the indicators that the administration proposed. Instead, the administration released an updated College Scorecard that "provides key measures of institutional performance in a clear, concise format designed to be easy to access on mobile devices" but stopped short of providing any actual rankings (NACUBO 2016).

As higher education in the United States and abroad enters the second generation of experience with ranking and universities continue to make sense of it, we should ask larger questions: Have rankings been a force for improvement in higher education? If so, for whom and in what ways? Does higher education ranking merely confuse status and prestige with quality, thereby reifying already elite universities? (Pusser and Marginson 2012). As ranking organizations continue their work, it is imperative that their influence continue to be the subject of rigorous empirical research.

References

Bastedo, Michael N., and Nicholas A. Bowman. 2010. *U.S. News & World Report* College Ranking: Modeling Institutional Effects on Organizational Reputation. *American Journal of Education* 116(2): 163–84.

Berdahl, Robert O. 1985. Strategy and Government: U.S. State Systems and Institutional Role and Mission. *International Journal of Institutional Management in Higher Education* 9(3): 301–7.

Berger, Peter L., and Thomas Luckmann. 1966. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Garden City, NY: Doubleday.

Bollen, Johan, Huina Mao, and Xiao-Jun Zeng. 2011. Twitter Mood Predicts the Stock Market. *Journal of Computational Science* 2(1): 1–8.

Bowman, Nicholas A., and Michael N. Bastedo. 2011. Anchoring Effects in World University Rankings: Exploring Biases in Reputation Scores. *Higher Education* 61(4): 431–44.

Brooks, Rachelle. 2005. Measuring University Quality. *Review of Higher Education* 29(1): 1–21.

Cooley, Alexander. 2015. The Emerging Politics of International Rankings and Ratings: A Framework for Analysis. In *Ranking the World: Grading States as a Tool of Global Governance*, edited by Alexander Cooley, and Jack Snyder, 1–38. Cambridge, UK: Cambridge University Press.

Cunha, Jesse M., and Trey Miller. 2014. Measuring Value-Added in Higher Education: Possibilities and Limitations in the Use of Administrative Data. *Economics of Education Review* 42: 64–77.

Diamond, Nancy, and Hugh D. Graham. 2000. How Should We Rate Research Universities? *Change: The Magazine of Higher Learning* 32(4): 20–33.

Dichev, Illa. 2001. News or Noise? Estimating the Noise in the *U.S. News* University Rankings. *Research in Higher Education* 42(3): 237–66.

Dill, David D., and Maarja Soo. 2005. Academic Quality, League Tables, and Public Policy: A Cross-National Analysis of University Ranking Systems. *Higher Education* 49(4): 495–533.

DiMaggio, Paul J., and Walter W. Powell. 1983. The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review* 48(2): 147–60.

Dranove, David, Daniel Kessler, Mark McClellan, and Mark Satterthwaite. 2003. Is More Information Better? The Effects of "Report Cards" on Health Care Providers. *Journal of Political Economy* 111(3): 555–88.

Eckerd, Adam, and Stephanie Moulton. 2011. Heterogeneous Roles and Heterogeneous Practices: Understanding the Adoption and Uses of Nonprofit Performance Evaluations. *American Journal of Evaluation* 32(1): 98–117.

Ehrenberg, Ronald G. 2002. *Tuition Rising: Why College Costs So Much*. Cambridge, MA: Harvard University Press.

Erkkilä, Tero. 2013. *Global University Rankings: Challenges for European Higher Education*. New York: Palgrave Macmillan.

Espeland, Wendy N. 1998. *The Struggle for Water: Politics, Rationality, and Identity in the American Southwest*. Chicago: University of Chicago Press.

Espeland, Wendy N., and Michael Sauder. 2007. Rankings and Reactivity: How Public Measures Recreate Social Worlds. *American Journal of Sociology* 113(1): 1–40.

Espeland, Wendy N., and Mitchell L. Stevens. 1998. Commensuration as a Social Process. *Annual Review of Sociology* 24: 313–43.

Frederickson, H. George. 2001. Getting Ranked. *Change: The Magazine of Higher Learning* 33(1): 48–55.

Frederickson, H. George., and Edmund C. Stazyk. 2010. Ranking U.S. Public Affairs Educational Programmes: Searching for Quality, Finding Equilibrium? In *Paradoxes of Modernization,* edited by Helen Margetts, Perri 6, and Christopher Hood, 63–80. Oxford, UK: Oxford University Press.

Fung, C.H., Y. Lim, S. Mattke, C. Damberg, and P.G. Shekelle. 2008. Systematic Review: The Evidence That Publishing Patient Care Performance Data Improves Quality of Care. *Annals of Internal Medicine* 148(2): 111–23.

Gnolek, Shari L., Vincenzo T. Falciano, and Ralph W. Kuncl. 2014. Modeling Change and Variation in U.S. *News & World Report* College Rankings: What Would It Really Take to Be in the Top 20? *Research in Higher Education* 55(8): 761–79.

Gormley, William T., and David.L. Weimer. 1999. *Organizational Report Cards*. Cambridge, MA: Harvard University Press.

Gower, John C. 1971. A General Coefficient of Similarity and Some of Its Properties. *Biometrics* 27(4): 857–71.

Graham, Hugh D., and Nancy Diamond. 1997. *The Rise of American Research Universities: Elites and Challenges in the Postwar Era*. Baltimore: Johns Hopkins University Press.

Grimmer, Justin. 2010. A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Release. *Political Analysis* 18(1): 1–35.

Hanushek, Eric A., and Margaret E. Raymond. 2004. Does School Accountability Lead to Improved Student Performance? *Journal of Policy Analysis and Management* 24(2): 297–327.

Hazelkorn, Ellen. 2011. *Rankings and the Reshaping of Higher Education: The Battle for World-Class Excellence*. New York: Palgrave Macmillan.

Hood, Christopher. 2012. Public Management by Numbers as a Performance-Enhancing Drug: Two Hypotheses. Special issue, *Public Administration Review* 72: S85–92.

Hood, Christopher, Ruth Dixon, and Craig Beeston. 2008. Rating the Rankings: Assessing International Rankings of Public Service Performance. *International Public Management Journal* 11(3): 298–328.

Iacobucci, Dawn. 2013. A Psychometric Assessment of the *Businessweek, U.S. News & World Report*, and *Financial Times* Rankings of Business Schools' MBA Programs. *Journal of Marketing Education* 35(3): 204–19.

Jones, Robert L. 2013. A Longitudinal Analysis of the *U.S. News* Law School Academic Reputation Scores between 1998 and 2013. *Florida State University Law Review* 40(4): 721–1027.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Kane, Thomas J., and Douglas O. Staiger. 2002. The Promise and Pitfalls of Using Imprecise School Accountability Measures. *Journal of Economic Perspectives* 16(4): 91–114.

Lavertu, Stéphane. 2015. We All Need Help: "Big Data" and the Mismeasure of Public Administration. *Public Administration Review.* Published electronically on August 18. doi:10.1111/puar.12436.

Lovett, Clara M. 2005. The Perils of Pursuing Prestige. 2005. *Chronicle of Higher Education,* January 21. 51(20): B20.

Luca, Michael, and Jonathan Smith. 2013. Salience in Quality Disclosure: Evidence from the U.S. News College Rankings. *Journal of Economics and Management Strategy* 22(1): 58–77.

March, James G., and Johan P. Olsen. 1989. *Rediscovering Institutions: The Organizational Basis of Politics*. New York: Free Press.

———. 1995. *Democratic Governance*. New York: Free Press.

March, James G., and Herbert A. Simon. 1958. *Organizations*. New York: Wiley.

Meredith, Marc. 2004. Why Do Universities Compete in the Ratings Game? An Empirical Analysis of the Effects of the *U.S. News & World Reports* College Rankings. *Research in Higher Education* 45(5): 443–61.

Merton, Robert K. 1968. *Social Theory and Social Structure*, Enlarged ed. New York: Free Press.

Morphew, Christopher, and Jeroen Huisman. 2002. Using Institutional Theory to Reframe Research on Academic Drift. *Higher Education in Europe* 27(4): 492–506.

Mukamel, Dana B., Simon F. Haeder, and David L. Weimer. 2014. Top-Down and Bottom-Up Approaches to Health Care Quality: The Impacts of Regulation and Report Cards. *Annual Review of Public Health* 35: 477–97.

Narins, Craig R., Ann M. Dozier, Frederick S. Ling, and Wojciech Zareba. 2005. The Influence of Public Reporting of Outcome Data on Medical Decision Making by Physicians. *Archives of Internal Medicine* 165(1): 83–87.

National Association of College and University Business Officers (NACUBO). 2016. Ed Launches New College Scorecard. September 21. http://www.nacubo.org/Business_and_Policy_Areas/Student_Financial_Services/Student_Financial_Services_News/ED_Launches_New_College_Scorecard.html [accessed June 24, 2016].

Obama, Barack. 2013. Remarks by the President on College Affordability. White House Office of the Press Secretary. August 22. https://www.whitehouse.gov/the-press-office/2013/08/22/remarks-president-college-affordability-buffalo-ny [accessed June 13, 2016].

Pusser, Brian, and Simon Marginson. 2012. The Elephant in the Room: Power, Politics, and Global Rankings in Higher Education. In *The Organization of Higher Education: Managing Colleges for a New Era*, edited by Michael Bastedo, 86–117. Baltimore: Johns Hopkins University Press.

Rothwell, Jonathan. 2015. Using Earnings Data to Rank Colleges: A Value-Added Approach Updated with College Scorecard Data. Report, Brookings Institution, October 29. http://www.brookings.edu/research/reports2/2015/10/29-earnings-data-college-scorecard-rothwell [accessed June 13, 2016].

Sanger, Mary Bryna. 2008. From Measurement to Management: Breaking through the Barriers to State and Local Performance. Special issue, *Public Administration Review* 68: S70–85.

Sauder, Michael, and Wendy N. Espeland. 2009. The Discipline of Rankings: Tight Coupling and Organizational Change. *American Sociological Review* 74(1): 63–82.

Sauder, Michael, and Ryon Lancaster. 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): 105–34.

Sloan, Margaret F. 2009. The Effects of Nonprofit Accountability Ratings on Donor Behavior. *Nonprofit and Voluntary Sector Quarterly* 38(2): 220–36.

Stake, Jeffrey Evans. 2006. Minority Admissions to Law School: More Trouble Ahead, and Two Solutions. *St. John's Law Review* 80: 301–21.

Stephens-Davidowitz, Seth. 2014. The Cost of Racial Animus on a Black Candidate: Evidence Using Google Search Data. *Journal of Public Economics* 118: 26–40.

Strathern, Marilyn. 1997. "Improving Ratings": Audit in the British University System. *European Review* 5(3): 305–21.

———. 2000. The Tyranny of Transparency. *British Educational Research Journal* 26(3): 309–21.

Tversky, Amos, and Daniel Kahneman. 1974. Judgment under Uncertainty: Heuristics and Biases. *Science,* New series, 185(4157): 1124–31.

Van Dyke, Nina. 2005. Twenty Years of University Report Cards. *Higher Education in Europe* 30(2): 103–25.

Van Thiel, Sandra, and Frans L. Leeuw. 2002. The Performance Paradox in the Public Sector. *Public Performance and Management Review* 25(3): 267–81.

Van Vught, Frans. 2008. Mission Diversity and Reputation in Higher Education. *Higher Education Policy* 21(2): 151–74.

Weber, Max.1968. *The Protestant Ethic and the Spirit of Capitalism: And Other Writings*. New York: Penguin, 2002

Webster, David S. 1981. Advantages and Disadvantages of Methods Assessing Quality. *Change: The Magazine of Higher Learning* 13(7): 20–24.

Zwick, Rebecca. 2007. College Admissions in Twenty-First-Century America: The Role of Grades, Tests, and Games of Chance. *Harvard Educational Review* 77(4): 419–29.