A Predictive Model For Benchmarking Academic Programs (pbap) Using U.S. News Ranking Data For Engineering Colleges Offering Graduate Programs

Lisa Gay Marie Chuck
University of Central Florida

A PREDICTIVE MODEL FOR BENCHMARKING ACADEMIC PROGRAMS (pBAP) USING *U.S. NEWS* RANKING DATA FOR ENGINEERING COLLEGES OFFERING GRADUATE PROGRAMS

by

LISA G.M. CHUCK
B.A. University of Central Florida, 1984
M.A. Columbia University/Teachers College, 1990

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Education
in the Department of Educational Research, Technology, and Leadership
in the College of Education
at the University of Central Florida
Orlando, Florida

Spring Term
2005

Major Professor: LeVester Tubbs

ABSTRACT

Improving national ranking is an increasingly important issue for university administrators. While research has been conducted on performance measures in higher education, research designs have lacked a predictive quality. Studies on the *U.S. News* college rankings have provided insight into the methodology; however, none of them have provided a model to predict what change in variable values would likely cause an institution to improve its standing in the rankings.

The purpose of this study was to develop a predictive model for benchmarking academic programs (pBAP) for engineering colleges. The 2005 *U.S. News* ranking data for graduate engineering programs were used to create a four-tier predictive model (pBAP). The pBAP model correctly classified 81.9% of the cases in their respective tier. To test the predictive accuracy of the pBAP model, the 2005 *U.S .News* data were entered into the pBAP variate developed using the 2004 *U.S. News* data. The model predicted that 88.9% of the institutions would remain in the same ranking tier in the 2005 *U.S. News* rankings (compared with 87.7% in the actual data), and 11.1% of the institutions would demonstrate tier movement (compared with an actual 12.3% movement in the actual data). The likelihood of improving an institution's standing in the rankings was greater when increasing the values of 3 of the 11 variables in the *U.S. News* model: peer assessment score, recruiter assessment score, and research expenditures.

This dissertation is dedicated to my parents, Eileen and Phillip Chuck;

my brother and sister, Andrew Chuck and Nicola Kelly;

and my love, William Massi III.

Their love and support motivated me to continue my studies

when I was overwhelmed with working full-time and completing my doctorate.

They never became upset during the numerous times

I needed solitude to work on my research.

ACKNOWLEDGMENTS

In the fall of 2001, I sat down with Dr. Carol Wilson and Dr. LeVester Tubbs for the first time to discuss entry into the Educational Leadership, Higher Education, doctoral program. Dr. Wilson guided me through the program until her retirement in 2003, and Dr. Tubbs continued on in her place as my advisor and my dissertation committee chair to the completion of my program this spring 2005. Dr. Doug Magann's rigorous standards for writing papers in his classes prepared me to write this dissertation. Dr. Stephen Sivo is one of the best teachers I have had; his passion for statistics is contagious and sparked my interest in statistics. Dr. Ruby Evans, with her eagle eyes, provided editorial feedback on this dissertation. In the spring of 2002, I met and was hired by Dr. Jamal Nayfeh, Associate Dean of Academic Affairs, College of Engineering and Computer Science. Dr. Nayfeh has supported me in many ways as his graduate assistant; as Assistant Director of Assessment, Accreditation, and Data Administration in the college; and as a member of my dissertation committee. To these members of my dissertation committee, I thank them for their encouragement, guidance, and support.

To my student assistants, my thanks; to Jason LeClair, for assisting with the tedious data pulling and setting up the Excel files; to Mary Dorsey, for helping with obtaining articles for my literature review; and to Manish Billa and Maciej Sobocinski, for developing the Web-interactive rank calculator. Special thanks to my friend, Kathy Niemczyk, for taking time out of her busy schedule to proofread this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

OVERVIEW OF THE STUDY

Introduction

College rankings have become a lucrative and competitive business (Pike, 2004; Walpole, 2003; Hunter, 1995; Dichev, 2001; McDonough, Lising Antonio, Walpole, & Xóchitl-Pérez, 1998; Webster, 1992a). It is common practice for universities not ranked in the top tier in national rankings to develop strategies for improving their standing (Arnone, 2003; Kleiner, 2004; Stecklow, 1995; Karl, 1999; Mallette, 1995; Dichev, 2001; Hunter, 1995; Thompson, 2000; Hossler, 2000; Ridley, Cuevas, & Matveev, 2001). Stakeholders of the university (prospective students, parents, legislators, and others) often consult college rankings before making a sizable investment in a college education (Dichev, 2001; Thompson, 2000; Stecklow, 1995; Wright, 1990–1991; Hunter, 1995; McDonough et al., 1998; Drew & Karpf, 1981; McDonough, 1994). The *U.S. News & World Report* (hereafter referred to as *U.S. News*) rankings are often consulted as the most popular college rating resource by the public, students, parents, legislators, and university administrators (McDonough et al., 1998; Dichev, 2001; Thompson, 2000; Karl, 1999; Stuart, 1995; Hossler, 2000; Ridley et al., 2001; Webster, 1992a). Despite criticisms of the *U.S. News* ranking methodology by educators (Pike, 2004; Dichev, 2001; Walpole, 2003; McDonough et al., 1998; Tsakalis & Palais, 2004; Thompson, 2000; Kleiner, 2004; Gottlieb, 1999; Vojak, Price, & Carnahan, 2003; Stuart, 1995; Mallette, 1995; Clarke, 2001; Ehrenberg, 2003; Webster, 1992a; El Khawas, 1992), educators will brag about their college's high rankings on their college Web sites and in their promotional materials (Kleiner, 2004; Thompson, 2000; Arnone, 2003; Hossler, 1998, 2000). The *U.S. News* rankings continue to influence public opinion (Stuart, 1995; Thompson, 2000; Monks & Ehrenberg, 1999); prospective students' choice of colleges (Walpole, 2003; Hunter, 1995; Hossler & Foley, 1995; Webster, 1992a); alumni and trustees perception of an institution's academic reputation (Hossler, 1998, 2000); and legislative and external funders' decisions (Arnone, 2003; Karl, 1999).

As a performance measurement tool, benchmarking has long been used by businesses in decision making to improve performance by comparing internal practices with best practices elsewhere (Reider, 2000; Karlöf & Östblom, 1993; Fombrun, 1996). Benchmarking has been used effectively in higher education research on tuition and fee decisions (Hubbell, Massa, & Lapovsky, 2002); strategic planning and tenure promotion procedures (Secor, 2002); distance education (Novak, 2002); policymaking of state higher education boards (Barak & Kniker, 2002); and college choice of high ability students from middle and upper-middle class families (Hossler & Foley, 1995; McDonough, 1994; McDonough et al., 1998; Walpole, 2003). Benchmarking is higher education's version of "comparison shopping"—a common American phenomenon (McGuire, 1995; McDonough et al., 1998; Rogers & Rogers, 1997; Hossler, 2000). College ranking publications are a form of benchmarking: they allow consumers to compare the "quality" of higher education institutions and to make an informed decision (Hossler, 1998; Webster, 1986). The popularity of the college rankings is not only due to the timeliness, accessibility, and easy-to-understand format of the rankings, but most of all to the fact that they provide cross-comparative information on academic reputation not available from universities (Walpole, 2003; Webster, 1992a, 1992b).

These terms will be used interchangeably throughout this paper: (a) indicators, independent variables, or variables; (b) tiers or groups; (c) *U.S. News & World Report* or *U.S. News*, (d) formula, algorithm, model, or methodology, and (e) cases or institutions .

Statement of the Problem and Purpose of the Study

Improving national ranking is an increasingly important issue for university administrators (Arnone, 2003; Kleiner, 2004; Stecklow, 1995; Karl, 1999; Mallette, 1995; Dichev, 2001; Hunter, 1995; Thompson, 2000; Hossler, 2000; Ridley et al., 2001). While research has been conducted on performance measures in higher education, research designs have lacked a predictive quality. According to Doerfel and Ruben (2002):

> As with business, higher education indicators have tended to be primarily historical, limited in predictive power, often incapable of alerting institutions to changes in time to respond, and lacking adequate attention to important but difficult-to-quantify dimensions. Ironically, the emphasis on easy-to-quantify, limited measures has, in a manner of speaking, "come to haunt" in the form of popularized college rating systems with which educators generally are frustrated and critical. But these are used consistently as the measures against which universities are evaluated by their constituents. (p. 20)

The purpose of this study was to develop a predictive model for benchmarking academic programs (pBAP) for engineering colleges offering graduate programs, based on the *U.S. News* college ranking methodology. This study was not concerned with the validity of the *U.S. News* methodology, but rather with the development of a predictive model within the existing *U.S. News* methodology. The pBAP model can be used to predict what changes in values would cause a college to move up or down a tier in ranking over time. A replicate study of the *U.S. News* methodology was also developed in this study to predict discrete changes in a college's ranking from year to year.

Conceptual Framework

The impact of college rankings on the higher education community has captured the attention of institutional researchers. Marsha Lichtenstein, Senior Institutional Researcher at the University of New Mexico, demonstrated an interactive Excel spreadsheet that she created in an attempt to reconstruct *U.S. News* rankings at the university level at the Association for Institutional Research (AIR) conference, May 2004). The interactive spreadsheet was created as a tool for university administrators to change values of the variables to determine how to move up their institutions in the rankings. Lichtenstein was unable to replicate the rankings with 100% accuracy due to missing ranking data; nevertheless, interest in the Excel template that she had created was high. The session was so well attended that some people sat on the floor and others overflowed into the hall. Moreover, Robert Morse, Director of Data Research, *U.S. News,* was in attendance at the session. In a later session conducted by *U.S. News,* Morse stated that he was encouraged to see researchers attempting to reconstruct the *U.S. News* ranking methodology to demystify how rankings are computed. The pBAP model developed in this study was also presented as a poster at the AIR conference and garnered a lot of interest. This study has also been accepted as a research presentation at the 2005 AIR Forum, May 29–June 1, 2005, San Diego, California.

While the *U.S. News* college rankings do not measure those "important but difficult-to-quantify dimensions" (Doerfel and Ruben, 2002, p. 20), these college rankings are widely read (McDonough et al., 1998; Dichev, 2001; Thompson, 2000; Karl, 1999; and Stuart, 1995). Published research on the behavior of the *U.S. News* methodology to rank universities was more prevalent than rankings available at the college and program level, perhaps because institutional researchers were more likely to conduct ranking studies, and their focus would be on the institution as a whole rather than the individual colleges within the institution. (Published research is defined as research readily available to the public regardless of the medium; for example, paper or electronic journal articles, research papers or articles available on the Internet through professional associations, research papers or articles available on the Internet in academic databases accessible through university libraries, etc.) Using *U.S. News* university-level data, researchers have studied:

1. the *U.S. News* college ranking methodology to examine the predictability of ranking changes (Dichev, 2001);
2. the relationship between a college's score on the National Survey of Student Engagement (NSSE) with its *U.S. News* ranking (Pike, 2004);
3. the relationship between the *U.S. News* rankings and the peer assessment scores of the National Research Council survey (Rogers & Rogers, 1997);
4. the role of *U.S. News* institutional rankings in determining freshmen's choice of colleges (McDonough et al., 1998; Walpole, 2003);
5. the validity and reliability of the *U.S. News* institutional ranking methodology (Thompson, 2000; Karl, 1999; Gottlieb, 1999; Wright, 1990–1991; Ehrenberg, 2003; Webster, 1992a; McGuire, 1995);

6. the effect of changes in the *U.S. News* methodology of five graduate discipline rankings and two university rankings over a six-year period (1995–2000) on interpreting shifts in academic quality (Clarke, 2001);
7. the frequency of an institution's tier movement in the *U.S. News* college rankings for four-year liberal arts colleges over a six-year period (1996–2001) (Ridley et al., 2001);
8. the impact of the rankings on institutional administrators (Hossler, 1998);
9. the impact of change in the rankings for the top-ranked national universities and liberal arts colleges on the number of students admitted and those that eventually enroll, average SAT scores, and tuition cost actually paid by students (Monks & Ehrenberg, 1999).

During the course of this study, the author found only three research studies on the *U.S. News* methodology to rank engineering colleges: (a) a study involving a control model to determine the most efficient way to raise an engineering college's standing in the *U.S. News* graduate rankings (Tsakalis & Palais, 2004); (b) a correlational study between the *U.S. News* college rankings for engineering graduate programs and the department (specialty) rankings (Vojak et al., 2003); and (c) a study on the effect of changes in the *U.S. News* specialty ranking methodology, including engineering, over a six-year period (1995–2000) on interpreting shifts in academic quality (Clarke, 2001). While this body of research provided insight into the *U.S. News* ranking methodology, none of them provided a model to predict what change in variable values would likely cause an institution to improve its standing in the rankings.

Most useful as a guide for this study was research by Ridley et al. (2001) on the frequency of transition between tiers for *U.S. News* college rankings for four-year liberal arts colleges over a six-year period (1996–2001). Their recommendations for future research, although directed at the institutional level, supported the methodology of this study for engineering colleges. First, replicate their study on tier movement for other categories of institutions; second, identify which categories of institutions were most likely to have tier movement; and third, perform discriminant analysis or logistic regression to determine which variables best discriminated between groups and use six-year averages to identify movers and non-movers (Ridley et al., 2001). The author of this study used the 2005 *U.S. News* ranking data for colleges of engineering offering graduate programs to create a four-tier predictive model (pBAP). The model can be used to predict what changes in values would cause a college to move up or down a tier in ranking over time. Classification discriminant analysis (DA) was deemed the appropriate statistical test to develop the pBAP model. DA identifies which variable means differentiate between predefined groups and predicts group membership of changed variable values or new cases (StatSoft Inc., 2004). The results of the DA test addressed the three recommended areas of research by Ridley et al. (2001) mentioned above for engineering colleges offering graduate programs ranked by *U.S. News*.

Research Questions

Research questions for this study were based on the *U.S. News* ranking methodology for engineering colleges offering graduate programs. There are 11 variables in the *U.S. News* methodology (see Table 1 in Chapter 3). Reputational scores are heavily weighted in the model; therefore, a hypothesis is that these scores will exert a strong influence on the overall ranking score. The research questions were:

1. Which *U.S. News* indicators contributed the most to the prediction model to discriminate between rankings or tiers?
2. What phenomena influenced *U.S. News* peer score assessment of academic reputation?
3. What was the predictive accuracy of the replicate *U.S. News* model and the pBAP model?

Data Sets of Interest

The predictive model developed in this study was based on the 2005 *U.S. News* data for engineering colleges offering graduate programs. The data set contained published rankings for 87 (of 167) engineering colleges. *U.S. News* declined a request to provide the full data set (personal communication, June 10, 2004). The 2005 *U.S. News* data were used to develop the pBAP model. To test the predictive accuracy of the pBAP model, the 2005 *U.S. News* data were entered into the pBAP variate developed using the 2004 *U.S. News* data, which contained data on 84 engineering colleges. Although a data set was available for the 2003 *U.S. News* rankings, it contained rankings for only 50 engineering colleges, which was an insufficient number of cases to create the four-tier pBAP model. Thus, these data were not used for the pBAP model. Instead, the 2003 *U.S. News* data set was used to examine the stability of the 20 top-ranked engineering colleges over a three-year period (2003–2005) and not in the predictive model. The data were collected by purchasing online premium access to the *U.S. News* Web site and were stored each year in an Excel file by the researcher. The *U.S. News* variables and methodology are listed in Table 1, which will be discussed in Chapter 3.

Operational Definitions

The definitions of the statistics used to create the predictive model in Chapters 3 and 4 are described below.

Z Score — standardized scores calculated to determine the relative distance of a data value (measured in standard deviations) from the mean of the group (SPSS Inc., 1999). The formula to calculate a *z* score is the variable value (x) minus the mean of the group ($\bar{x}$) divided by the standard deviation (s) of the group (Equation 1).

$$z = \frac{x - \bar{x}}{s} \quad (1)$$

Classification Discriminant Analysis (DA) — identifies the independent variable means that differentiate between predefined groups and predicts group membership of new cases or changed variable values (StatSoft Inc., 2004). A variate, a linear equation of weighted variables, is summed to calculate a discriminant score for each case to determine group membership (Hair, Anderson, Tatham, & Black, 1998). The independent variables are metric, and the dependent variables are categorical.

Independent Variables (Indicators) — *U.S. News* 11 indicators for engineering colleges offering graduate programs used in this study.

Dependent Variables — Division of cases into four groups (Tiers 1, 2, 3, and 4) in this study.

Cross-Validated Classification Matrix — compares the predictive accuracy of the predicted group membership with the actual group membership. The sample is divided into an analysis sample and a holdout sample, and each sample is used to validate the predictive accuracy of the discriminant function(s) to avoid overclassifying cases into the groups (Hair et al., 1998). SPSS leaves out one case at a time (the case being classified) for cross-validation (SPSS Inc., 1999), by using a test called "jackknifing" (Hair et al., 1998).

Proportional Chance Criterion — determines if the cross-validated classification accuracy in the model predicts cases into groups better than by chance occurrence alone (Hair et al., 1998).

Press' Q Statistic — determines if the discriminatory power (cross-validated classification accuracy) of the model is statistically significant relative to classification accuracy by chance occurrence alone (Hair et al., 1998).

Box's M Statistic — determines if the covariance matrices of the variables are equal across groups (Statsoft Inc., 2004).

Mahalanobis D^2 — determines if a case is an outlier by comparing its Mahalanobis D^2 value with a critical value for a probability of .01 with the degrees of freedom equal to the number of independent variables used to compute the Mahalanobis D^2 (Hair et al., 1998; Schwab, 2003).

Canonical Correlation — determines the extent to which each discriminant function explains the variation in group membership by calculating the relationship between the discriminant scores and the groups (SPSS Inc., 1999). Discriminant analysis derives g-1 functions based on the number of groups into which cases could be classified

(Statsoft, 2004). There were four groups (tiers) in this study; therefore, three discriminant functions were derived. Only functions that are statistically significant should be interpreted; functions that are not statistically significant are ignored (Statsoft Inc., 2004).

Canonical Discriminant Function Coefficients — Unstandardized coefficients provide the independent variable weights and a constant for the discriminant variate, which when summed produce the discriminant score (Statsoft Inc., 2004). Standardized coefficients explain the unique contribution of each independent variable to the discrimination between groups while controlling for the other variables. Larger variable weights signify greater contribution to the classification of cases into groups.

Centroid — used to calculate the cutoff discriminant score to predict group membership (Statsoft Inc., 2004). The centroid is the mean of the discriminant scores of all the cases within each group; for groups of equal size, the optimal cutting score is the halfway point between two group centroids (Hair et al., 1998). There were four groups in this study; therefore, four group centroids (or means) were produced.

Structure Matrix Coefficients — determine the contribution of each independent variable as influenced by other independent variables with the discriminant function (Statsoft Inc., 2004). Larger coefficients exert a stronger influence in explaining the variance in the function. Structure coefficients should be used to assign labels to the discriminant function (Statsoft Inc., 2004).

Assumptions of Classification Discriminant Analysis

Hair et al. (1998) cited two key assumptions for classification discriminant analysis (DA): (a) multivariate normality of the independent (discriminating) variables, and (b) equality of the covariance matrices for each group. DA is sensitive to departures from multivariate normality and equal covariances for the groups, but can be ignored and the results interpreted if certain conditions are met (Hair et al., 1998; Statsoft Inc., 2004). Violation of these assumptions is considered minimal if groups are approximately equal in size, and "if the largest group size divided by the smallest group size is less than 1.5" (Hair et al., 1998, p. 348). Other steps should be taken to improve the homogeneity violence. Independent variables not meeting normality assumptions can be transformed by computing the log 10, square root, inverse, or square root values of the variables (Schwab, 2003). If the transformed variables meet the normality assumption, then they can be substituted in the data set to determine if the Box's *M* statistic has improved as well as the cross-validated classification accuracy. Outlier variables can also be examined and kept or removed as appropriate (Hair et al., 1998).

Methodology

Classification discriminant analysis (DA) was deemed the appropriate statistical test to develop the pBAP model. DA identifies which variable means differentiate between predefined groups and predicts group membership of changed variable values or new cases (StatSoft Inc., 2004). A variate, a linear equation of weighted variables, is summed to calculate a discriminant score for each case to distinguish group membership (Hair et al., 1998). Statistical Package for the Social Sciences (SPSS) software was used to analyze the data. This study also included the replication of the *U.S. News* ranking methodology for engineering colleges offering graduate programs to predict discrete changes in ranking. Together, the pBAP tier model and the *U.S. News* replication model provide a tool to predict tier ranking (longitudinal goal) and discrete ranking (short-term goal).

Significance of the Study

Historically, the *U.S. News* college rankings have impacted public perception of an institution's academic reputation. Unless universities, colleges within universities, and individual disciplines can agree on a common set of measures to rate educational quality, then the *U.S. News* college rankings are likely to continue influencing public opinion. Although the pBAP model in this study used the *U.S. News* ranking data specifically for graduate engineering colleges, the methodology can be adapted for other disciplines. The methodology in this study can also be adapted to develop a model at the university level. As part of this study, a Web-interactive ranking calculator was developed to predict tier movement (completed) based on the pBAP model and discrete movement (in progress) based on the *U.S. News* replicate model, as a tool for engineering college administrators to predict ranking. The Web rank calculator will be demonstrated to institutional researchers and university administrators at the 2005 Association for Institutional Research (AIR) conference in the research presentation category.

Limitations and Delimitations

Three limitations derive from the current model to predict the following year's rankings: (1) assumption that data for all other institutions other than the selected institution remain constant when predicting future *U.S. News* rankings; (2) potential changes to the *U.S. News* ranking algorithm; and (3) missing data from *U.S. News* non-ranked (not published) institutions. According to Robert Morse, Director of Data Research, *U.S. News*, changes are made to its ranking algorithms periodically in response to suggestions for improvement from the higher education community (Morse & Gilbert, 1995). *U.S. News* (2003) cautions that rankings are not necessarily comparable from year to year; therefore, the data on the colleges and not necessarily the rankings should be compared for year-to-year comparisons. A study by Clarke (2001) compared the 1995 and 2000 *U.S. News* graduate rankings for engineering and found a .88 correlation, taking into account changes to the methodology, thereby demonstrating that changes to the engineering methodology did not have a significant impact on the rankings. Based on the findings of Clarke's (2001) study, future changes by *U.S. News* to the engineering ranking methodology should have minimal impact on the rankings from the previous year. As a precaution, however, the discriminant analysis procedure should be rerun with the current data each year to determine the new weights for the pBAP model for the following year's predictions.

Organization of the Study

This dissertation is organized into five chapters. Chapter 1 provides an overview of the study. Chapter 2 summarizes the impact of the *U.S. News* college rankings on reputational ratings, the global marketplace, external funding, and students' college choices. Chapter 2 also covers background information on the *U.S. News* college rankings. Chapter 3 describes the methodological approach used to develop the predictive model for benchmarking academic programs (pBAP) and to replicate the *U.S.*

News ranking methodology for engineering colleges. Chapter 4 presents the results of the data analyses and the pBAP model. Chapter 5 presents findings and recommendations from the study and implications for future research.

CHAPTER 2

REVIEW OF LITERATURE

Introduction

The attempt to rank colleges began in 1870 with a report of institutional data published by the United States Bureau of Education, which gradually expanded to include more data and a ranking of institutions (Stuart, 1995). Although not originally intended as a comparative resource for academic quality rankings, James Cattell's 1906 publication, *American Men of Science*, served as a useful reference for that purpose (Webster, 1986). Cattell later published his influential academic quality rankings of American research universities in 1910 based on the research reputation of faculty (Stuart, 1995; Webster, 1986, 1992b). It was about this same time that professional accrediting associations and state accrediting boards were founded, but in some cases, reported only if an institution or program was accredited or not (Stuart, 1995). While other publications on academic quality rankings followed, comparative data on academic quality ratings remained a scarce commodity, and those that were available were aimed at academicians and legislators rather than the general public (Stuart, 1995).

It was not until the 1970s and 1980s, when intense public scrutiny and declining college enrollment forced colleges to adopt business marketing and advertising strategies, that college ranking and guidebook publications became a booming market available to the general public (Hunter, 1995). College guidebooks provided comparative information on programs but not rankings of academic quality (Webster, 1986). Considered the "gold standard of the [college] ranking business" (Ehrenberg, 2003), the *U.S. News & World Report* college rankings made their debut in 1983, the same year that the influential and highly critical publication, *Nation at Risk,* was released by the National Commission on Excellence in Education (Clarke, 2001). The report warned of the impending jeopardy of America's position as a superpower in the world global marketplace if the American educational system was not reformed to be competitive with other leading nations (Bracey, 2003).

In a study of 16 of the top 25 national universities and the top 25 national liberal arts colleges in the 1998 *U.S. News* rankings, Monks and Ehrenberg (1999) examined the impact of *U.S. News* rankings on students' college choices. They found that institutions that experienced a decrease in the rankings were likely to attract fewer students, thereby forcing the institution to admit more applicants (increasing their selectivity rate), and fewer of those students tended to enroll (decreasing their yield) (Monks & Ehrenberg, 1999). Despite criticisms of the *U.S. News* ranking methodology by educators (Pike, 2004; Dichev, 2001; Walpole, 2003; McDonough et al., 1998; Tsakalis & Palais, 2004; Thompson, 2000; Kleiner, 2004; Gottlieb, 1999; Vojak et al., 2003; Stuart, 1995; Mallette, 1995; Clarke, 2001; Ehrenberg, 2003; Webster, 1992a; El Khawas, 1992), according to Ehrenberg (2003), the *U.S. News* college rankings:

> . . .are probably more symptomatic of the increasingly competitive environment in which academic institutions find themselves than its underlying cause. . . [while] the *USNWR* ranking methodology provides incentives for institutions to take actions that are not always socially desirable, the methodology does not penalize institutions for cooperating in ways that improve the education they are providing for their students or for increasing the efficiency of their operations. (p. 158)

The Role of Higher Education in the Global Market

Higher education is increasingly viewed as a commodity to be traded in the global market (Altbach, 2001; Breton, 2002; The Futures Project, 2000). The U.S. Department of Commerce described education as the fifth largest service sector export in America (International Institute of Education, 2001). The World Bank (1994) reported that there was a correlation between the number of students enrolled in institutions of higher education and economic development. Henry, Lingard, Rizvi, & Taylor (2001) reported that the average population enrollment in higher education for high-income countries was 51%; in middle-income countries, 21%; and in low-income countries, 6%.

Universities are social institutions, and as such, are considered a public good serving the needs of a society on its path to economic advancement (Altbach, 2001). Access to knowledge provides social mobility for the individual and economic advantage for nations (Neave, retrieved January 19, 2003). According to Shea (2004a), bachelor's degree graduates in the U.S. earned an average of 89% more than high school graduates in 2001. About two million students around the world study overseas annually, a large portion of whom enroll in U.S. institutions (Dillon, 2004). During the 2000–2001 academic year, 547,867 international students were enrolled in U.S. colleges and universities and contributed $11 billion to the U.S. economy (International Institute of Education, 2001), now up to $13 billion according to Dillon (2004).

Proponents for the globalization of higher education argue that access to lifelong learning should be available to all (Rossman, 2002), while opponents take the stance that developing countries will not be able to compete in the global higher education marketplace (Altbach, 2001). UNESCO (as cited in Rossman 2002) has expressed lofty goals for higher education:

> Higher education should reinforce its role of service to society, especially its activities aimed at eliminating poverty, intolerance, violence, illiteracy, hunger, environmental degradation and disease, mainly through an interdisciplinary and transdisciplinary approach in the analysis of problems and issues. Ultimately, higher education should aim at the creation of a new society — non-violent and non-exploitive — consisting of highly cultivated, motivated and integrated individuals, inspired by love for humanity and guided by wisdom. (1.P.7)

Since 1972, the World Bank has spent in excess of 30 billion dollars to reduce poverty in developing countries, and the only solution to increasing access to education was to use technology, namely the Internet (Rossman, 2002). The annual expenditure per child in developed nations was $4,000–$5,000 annually, while in developing nations, it was a mere $150–$250 per child (Rossman, 2002). According to data from the College Board, 50% of the people in the United States and Canada are connected to the Internet, compared to only 2% in Asia, Latin America, the Middle East, and Africa (The Futures Project, 2000). While there has been an increase in enrollment of international students worldwide, enrollment has been stronger in the developed countries rather than emerging countries. In 1997, for example, there was one doctoral student per 5,000 inhabitants in developed countries; one per 70,000 in Brazil; one per 140,000 in Chile; and one per 770,000 in Colombia (Breton, 2002).

Globalization of higher education has spawned distance learning courses and created the virtual university (Salmi, 1999). Where students previously chose from 3,600 traditional institutions of higher education, they can now choose from 5,000 institutions (Newman & Couturier, 2001). Dominant market players are emerging in the global market of online distance learning. The now defunct Universitas 21, a joint venture of 18 leading global universities, was formed in May 2000 to compete with the University of Phoenix (The Futures Project, 2000). Other examples of distance-learning initiatives reported in The Futures Project (2000) included the following: New York University offered online courses to students from 35 countries; Tajikistan established an English-language distance university; the British Open University had 43 branches worldwide; and Australia's Monash University had two campuses in Malaysia and South Africa and offered 108 distance programs to students from 80 countries.

The European Union (EU) is also considering an "open door research area" with free mobility of researchers to strengthen research collaboration (Ince, 2004a). Emerging countries, such as India and China, have invested heavily in improving their higher education infrastructure to be competitive with developed countries and to combat "brain drain." Similarly, China has made it a national priority to transform "100 universities into world-class research institutions" (Dillon, 2004, ¶3). For example, the Indian Institute of Technology in India is ranked number 11 of the top 40 universities in the East (Ince, 2004b). India and China were the top exporting countries of college students during the academic year 2000–2001: 59,939 Chinese students and 54,664 Indian students were enrolled in colleges overseas (International Institute of Education, 2001). With improvements in the higher education infrastructure in China and India, a significant number of these college students may eventually choose to study in their home countries or nearby countries because of the affordability and convenience (Dillon, 2004).

College national rankings have become a lucrative business (Pike, 2004; Walpole, 2004; Hunter, 1995; Dichev, 2001; McDonough et al., 1998; Webster, 1992a) because they play an important role in American society. College global rankings are set to follow suit because they play an important role in the global market, as competition for the best and brightest students intensifies. According to O'Leary (2004), "Higher education has become so international that it is no longer enough for the leading universities to know that they are ahead of the pack in their own country" (p .2). For the moment, U.S. universities still claim the top spots; they appeared in 7 of the top 10 spots and in 14 of the top 25 spots in the ranking of the world's top 200 universities compiled by the Shanghai Jiao Tong University in China (O'Leary, 2004). Today, universities need to be ranked in the top 25 in the world to remain competitive.

The Cost of Higher Education

In the highly competitive student recruitment environment, college admissions officers eagerly tout their national standing in popular college ranking publications to students and parents (McGuire, 1995). Based on personal communication from John Katzman, CEO of the *Princeton Review,* McDonough et al. (1998) reported that students and parents spent an estimated $400 million annually on college preparation products, or an average of $250 per student. A college education is a substantial investment for students and legislators. "The cost of going to college catapulted up to the third place in the list of Americans' biggest worries during the mid-1990s" (Stanfield, 1998, p. 135). Morse and Gilbert (1995) estimated that for some students, a college education could cost upwards of $120,000. In 2002, institutions of higher education spent an average of $19, 220 per student, and states spent $66 million in subsidies to institutions of higher education (Gorman, 2003). The average tuition for students in four-year public colleges in 2002 was $4,081 (Gorman, 2003) and $5,132 in 2003–2004 (College Board, 2004). However, in 2003–2004, the average out-of-pocket tuition cost to students attending four-year public institutions was $1,300, after having excluded grant aid and education tax benefits (College Board, 2004). On average, student tuition payment to four-year public colleges covered 25% (calculated from the amounts above: $1,300 student out-of-pocket tuition divided by $5,132 student total tuition for 2003–2004) of the tuition cost in public colleges and universities, with the rest subsidized by the state. Students are increasingly bearing a larger portion of tuition costs because of rising tuition costs, increasing reliance on student loans, and decreased state funding for public institutions. In 1982, 84% of tuition costs at public universities was covered by Pell grants, whereas in 2002, it had declined to 42% (Shea, 2004a).

Deciding which college to attend is one of life's major choices students and parents must make. It requires a substantial commitment in both time and money (McGuire, 1995; Hossler & Foley, 1995). Yet, universities are often reluctant to publish or release certain information that could help students and parents make an informed decision because such information could potentially be harmful to the university's reputation (Hunter, 1995; Thompson, 2000; Morse & Gilbert, 1995; Wright, 1990–1991; McDonough, 1994). Assessing the quality of universities, colleges within universities, and individual programs is a complex task. Institutions have differing missions and defy standardization of a common definition of what constitutes quality (McGuire, 1995; Stuart, 1995). College rankings are another tool that can be consulted to demystify the evaluation of the complex organization that is the university (Morse & Gilbert, 1995; Hossler & Foley, 1995).

Legislators are wrestling with the issue of deciding whether to allocate funds to flagship universities or to "wannabe" regional universities (Arnone, 2003). The amount of money that is expended on education by legislators is substantial. Drucker (1969) reported that by the late 1960s, the American government spent more on education than all other non-defense community services combined. Education expenditures amounted to $70 to $75 billion annually, $50 billion of which was spent by school and university systems (public and private), with the remaining amount spent by industry, government, and the armed forces for schooling and training. More recently, of the approximately 3,000 postsecondary institutions in the United States, only about 100 received federal government science grants in excess of $50 million annually (Altbach, 1998). In 2002, the U.S. government spent $289 billion on U.S. institutions of higher education (Gorman, 2003), over six times the cost 30–40 years ago. Muffo (2003) reported that in 2000, the Harvard endowment was valued at around $19 billion and the public University of Texas system in excess of $10 billion. Endowments are over and above the billions allocated from federal and state governments to education. The competition for funding is fierce because the stakes are high.

Accountability

The public thinks that universities and colleges charge more for tuition than they actually do (Stanfield, 1998). In response to taxpayers' concerns relating to the allocation and expenditure of their tax dollars, government is demanding higher accountability of federal and state funding to educational activities (Griffin, 1999; Barak & Kniker, 2002; Altbach, 1998). The situation is becoming gloomier. For the fiscal year

2002–2003, half of the states in the U.S. had reduced or were planning to reduce appropriations for higher education (Arnone, Hebel, & Schmidt, 2003). In an effort to address a predicted $400 billion federal budget deficit, President Bush's proposal for cutting the budget for the 2006 fiscal year included reducing the budgets for research and student financial aid programs (Brainard & Burd, 2004). Bush's proposal would reduce by 2% the budgets of the National Institutes of Health (NIH) to $28 billion and the National Science Foundation (NSF) to $5.6 billion, and it would cut student financial aid by about $325 million. NIH and NSF are the largest funders for university researchers. Consequently, universities have turned to other activities to generate funding. Since legislation to allow universities to benefit from federally funded research was passed in 1980, 3,870 companies have been formed to create products from university-held patents and licenses (Shea, 2004a). Aggressive fund-raising by universities has increased private donations by 159% since 1980 (Shea, 2004a).

Higher education's reliance on government support is not unusual; in like manner, international institutions rely heavily on government support (The World Bank, 1994). Because globalization is so rapid and dispersed, the challenge is to ensure that programs offered by the global university comply with appropriate standards (The Futures Project, 2000). To address the issue of standardization of comparative programs globally, an independent quality assurance agency was established in Hong Kong in 1990, the Hong Kong Council of Academic Accreditation (HKCAA), with a membership of over 100 members from 50 countries (The Futures Project, 2000). There have been attempts to compile international college rankings. For example, Shanghai Jiao Tong University in China compiled a ranking of the world's top 200 universities using criteria such as peer review rating score, international faculty (on campus) score, international student (on campus) score, the number of faculty per student score, and the number of citations per faculty score (O'Leary, 2004).

According to Karlöf & Östblom (1993), "the purpose of all organized activity is TO CREATE VALUE WHICH IS GREATER THAN THE COST OF CREATING IT" (p. 3). College rankings, a form of benchmarking, can be a useful tool not only to diagnose the health of the organization on predetermined factors but also to meet compliance requirements. Reputation is increasingly being used as a tool for assessing companies (Fombrun, 1996).

Access to Higher Education and Student College Choice

A brief summary of the history of student college choice is presented in this paragraph, referenced from Frederick Rudolph's (1990) classic text, *The American College & University: A History*. Rudolph provided a comprehensive history of higher education from its beginning in colonial America. The purpose of the American colonial college was to preserve a civilization of cultured men in the wilderness of the New World in America. The colonial college was the medium through which the sons of the wealthy were educated to become clergy and lettered men to provide leadership to the nascent American culture. Beginning with the founding of the first American college in 1636, Harvard University, education was primarily a private concern, but not without the active but limited support of government. Over the next two hundred years, the purpose of education shifted from a private concern to a public concern; education now had a value to the masses as the vehicle to social and economic advancement based on ability and regardless of social standing and wealth or lack thereof. The Progressive Movement, approximately the period between the Spanish-American War and World War I, saw the rise of the middle class primarily in response to the industrialization of the agrarian economy, urbanization, and increased immigration. Colleges reached out to the farmers to provide relevant training to improve the efficiency of crop yield and animal husbandry. The federal government provided funding through the Morrill Land Grant Act of 1862, leading to the founding of the first research centers and public Agricultural and Mechanical (A&M) colleges. Farmers no longer needed their sons to work on the farm; they now had machinery and models of efficient farming that allowed them to work smarter not harder. Farmers turned to the colleges as a medium for their sons to join the ranks of the elite.

Other major federal policy initiatives followed to increase access to education for the masses. The 1890 Morrill Act provided for annual appropriations to the states if college admission policies did not discriminate based on race or had separate-but-equal facilities, and the GI bills of 1924 and subsequent versions of the bill provided education benefits for servicemen and their families (Rudolph, 1990). More recently, two influential federal policies enacted to redress discrimination were Title VI of the Civil Rights Act of 1964 and Affirmative Action in 1965 (Brunner, retrieved February 6, 2005). By 2000, 12% more women than men were enrolled on college campuses; 55% more African Americans were enrolled in 2001, up from 45% in 1972; and 52% more Hispanics were enrolled in 2001, up from 45% in 1972 (Shea, 2004a).

In the twentieth century, the value of a college education has increased sharply since the advent of the knowledge-based society (Breton, 2002). The disparity in salaries between high school graduates and college graduates has continued to grow (Stanfield, 1998). Deciding to attend college is one of life's major decisions that students and parents confront and which requires a substantial commitment in both time and money (McGuire, 1995; Hossler & Foley, 1995). College rankings are most likely to be consulted by high ability students from the middle and upper-middle classes (Hossler & Foley, 1995; McDonough, 1994; McDonough et al., 1998; Walpole, 2003). In fact, a study by McDonough et al. (1998) revealed that students who were most likely to consult college rankings were Asian American from affluent families with college-educated parents. They also tended to ask their high school teachers for advice; receive high grades; and have aspirations of doctoral, medical, or law degrees. In addition, they usually attended colleges outside of their hometowns; they were self-motivated, confident about their academic skills, and had high expectations of a satisfactory college experience; and they were more likely to apply to more universities and target private universities.

In the *U.S. News & World Report* special edition magazine, "America's Best Colleges: Exclusive 2005 Rankings," Ewers (2004) recounted the tale of a highly motivated high school student, Meaghan, and the stress and anxiety of her attempts to be admitted to the college of her choice. And the college admission process will become increasingly competitive as more students graduate and continue on to college — 3.3 million high school graduates by the year 2009 according to census data (Ewers, 2004). Moll and Wright (1998) attributed this college admission anxiety in students and parents to the influence of the media, "Reporters stalking and reporting on a handful of highly selective colleges have created serious anxiety that has incrementally escalated at the family dinner table and in high school corridors" (p. 149). There are articles, however, that provide a more balanced picture, such as the one by Sara Sklaroff (2004) in which she advised prospective college students ". . .that the right school for you is not necessarily the 'best' one you can get into by any measures other than your own desires. . . . In the end, it is you, and not the school you choose, who will make your college years a success" (p. 23). Nevertheless, the perception of a college's academic reputation can influence college choice. McDonough et al. (1998) found that of the students who consulted college rankings and believed them to be very important in college choice, 70% believed that academic reputation influenced selection of their graduate school of choice.

While the media has perpetuated the image of intense competition in the college admissions process by focusing on private colleges and universities (Stanfield, 1998), admission to public colleges and universities has remained accessible (Moll & Wright, 1998). In the mid-1990s, over half (53%) of U.S. colleges and universities admitted 75%–99% of their freshman applicants on an average, and 11% had open admissions (Moll & Wright, 1998). Access to higher education in America today has significantly improved from colonial America.

U.S. News & World Report College Rankings

Background

Since its inception in 1983, the *U.S. News & World Report* college rankings have become the most influential (McDonough et al., 1998; Dichev, 2001; Thompson, 2000; Karl, 1999; Stuart, 1995), timely, and accessible resource for information on colleges (Walpole, 2003). While *Money* magazine has published college rankings, their college rankings have not reached the popularity as those of *U.S. News*. *Money* does not publish the rankings as often as *U.S. News*; it does not publish the rankings in its main magazine with a circulation of more than 1.9 million, but rather in a separate publication with a circulation of about 300,000; and it does not dedicate many pages and prominent displays to its rankings (Webster, 1992a). On the other hand, examination of national bookstands and the *U.S. News* Web site showed that *U.S. News* rankings are widely disseminated, published annually or frequently, easily accessible online, and available in a variety of hard-copy publications (*U.S. News,* 2003, 2005).

The *U.S. News* college rankings began with a survey of university presidents to rank undergraduate programs of the five best peer colleges; these reputational rankings were first published in 1983 in the *U.S. News & World Report* magazine (Thompson, 2000). The success of the college ranking publication spawned additional publications from *U.S. News*: in 1987, professional school rankings at the graduate level in business, engineering, law, and medicine; in 1992, the social sciences (Webster, 1992a); in 1990, *America's Best Colleges* guidebook (undergraduate programs); and in 1992, *America's Best Graduate Schools* (Stuart, 1995). Further examination of the *U.S. News* Web site and national bookstores and newsstands showed that not only do these publications still exist today, but also college rankings for graduate programs in fine arts, education, health disciplines, library science, public affairs, and the sciences (*U.S. News,* 2003, 2005). In fact, their rankings are so popular that *U.S. News* has compiled data and published rankings on the best stocks, mutual funds, and hospitals in the U.S. (Webster, 1992a). The *U.S. News* college guidebooks alone have a circulation of 600,000 and remain on national newsstands for one year (Jim Robben, EMI Network, personal communication, April 15, 2004). An estimated 3.2 million copies of *U.S. News* college ranking issues are sold annually with a readership of 11 million people (Dichev, 2001). In addition to the college rankings, the guidebooks and the special edition magazine publish informative articles on the colleges to help prospective students select the best colleges (*U.S. News*, 2003). The *U.S. News* college ranking publications also allow university administrators, parents, students, and legislators to compare (i.e., benchmark) institutions across the nation on a variety of indicators.

Readers of the college ranking publications should be made aware that ranking is not a precise art (Roose & Andersen, 1970). *U.S. News* has been criticized for creating a college ranking algorithm that favors the Ivy Leagues and prestigious institutions, thereby predetermining the outcome of top-ranked schools (Thompson, 2000; Gottlieb, 1999). According to Webster (1992a), *U.S. News* based its first three rankings entirely on academic reputational scores, but later integrated objective measures into selected ranking methodologies (Webster, 1992a). The origin of the ranking algorithm is attributed to Mel Elfin, a former journalist with *Newsweek* hired by *U.S. News,* who in an interview with Thompson (2000) was quoted as saying, “When you’re picking the most valuable player in baseball and a utility player hitting .220 comes up as the MVP, it’s not right.” This analogy to baseball is the justification of tinkering with the original algorithm so that prestigious institutions came out on top (Thompson, 2000); in other words, other institutions are benchmarked against the prestigious institutions. As part of the research for Thompson’s article (2000), the *Washington Monthly* posted the results of a report commissioned by *U.S. News* in 1997 from the National Opinion Research Center (NORC) to assess the *U.S. News* ranking methodology. The NORC report identified poor justification of the weights for the indicators in the algorithm as one of its major criticisms (NORC, 1997, as posted by the *Washington Monthly*). *U.S. News'* response by Peter Cary, Special Projects Editor, and posted in the *Washington Monthly* (2000) provided examples of changes made to the *U.S. News'* methodology that were based on recommendations by the NORC report. Cary's article echoed an earlier article published in 1995 by the long-standing *U.S. News* Director of Data Research, Robert Morse, with Jersey Gilbert, formerly with *Money* magazine, in which Morse and Gilbert (1995) emphasized *U.S. News'* willingness to continue soliciting input on the ranking methodology from the academic community. Mel Elfin (1992) supported the input of the higher education community:

> …*U.S. News* takes very seriously its growing role as a prime purveyor to the consumer of information about what academicians lovingly refer to as "the enterprise"….how hard *U.S. News* works both to "get it right" and to improve our journalistic product….how eager we are to work with the higher education community at all levels.…contrary to those who carp, cavil, and criticize our "disservice to higher education," we serve to reform an enterprise that by its own admission has lost a good deal of the public trust. (p. 7)

Methodology for Engineering Graduate Programs

U.S. News began ranking engineering graduate programs in 1987 (Stuart, 1995; Webster, 1992a; Clarke, 2001). The *U.S. News* ranking model for colleges of engineering offering graduate programs was developed by Robert Morse with input from "various academics in the field" (Samuel Flanigan, Deputy Director of Data Research, *U.S. News,* personal communication, November 23, 2004). This paper, however, was not concerned with the validity of the *U.S. News* college ranking methodology. Information on the history and the controversy surrounding the *U.S. News* college rankings was included in the section above to provide an understanding of the ranking methodology. If the reader accepts that the rankings are biased toward the prestigious institutions, then the measures that are most likely to influence the predictive (pBAP) model in this study are the ones that distinguish the existing top-ranked institutions from the other institutions.

The *U.S. News* graduate engineering methodology uses *z* scores to determine overall rankings. The overall ranking score is the summation of the *z* score of 11 indicators, each of which is weighted. (See Table 1 in Chapter 3 for a full description of the methodology.) Academic reputation ranking scores within the *U.S. News* model for engineering colleges offering graduate programs are collected by surveys sent by *U.S. News* to deans and deans of graduate studies of engineering colleges (weighted at .25) and industry recruiters (weighted at .15), for an overall weight of .40 in the ranking algorithm (*U.S. News,* 2003). Therefore, there was a high probability that peer and recruiter scores would have the most impact on ranking in the prediction model (pBAP) developed in this study. If this hypothesis were accepted, then it would be important to determine the phenomena that influence rater bias in assessing academic quality (where quality is equated with academic reputation).

Research studies have suggested that a "halo effect" exists between peer assessment ratings of program quality with (a) program size (National Research Council, 2004; Astin, 1992), and (b) the reputation of a few well-known faculty (Stuart, 1995; Rogers & Rogers, 1997; Fombrun, 1996) also known as a "star effect" (National Research Council, 2004). If these phenomena held true, then there would be a high correlation between peer assessment score with (a) the number of full-time graduate students (halo effect), and (b) the number of full-time engineering faculty who were members of the prestigious National Academy of Engineering (star effect), all of which are indicators in the *U.S. News* engineering methodology. The profile of a successful engineering college could then be described as one that already had a strong academic reputation (usually established, elite universities), high graduate student enrollment, and many faculty with membership in the National Academy of Engineering, for example. Thus if institutions not ranked in the top tier had dreams of being ranked in the top tier, they would have to improve in these areas.

U.S. News' response to Gottlieb's 1999 article posted online by *Slate* indirectly supported the claim that the *U.S. News* rankings were biased toward the prestigious institutions; Peter Cary, *U.S. News* Special Projects Editor, and Brian Duffy, *U.S. News* Executive Editor (1999), wrote, "Over the past 10 years (1991–2000), the top national universities in the *U.S. News* rankings have remained remarkably consistent." Dichev (2001) examined the *U.S. News* rankings for the top 25 national universities and national liberal arts colleges over a 12-year period (1987–1998) and observed that "a striking feature that emerges from this examination is that the composition of the Top 25 schools is remarkably stable over time" (p. 241). Clarke (2001) investigated the stability of the *U.S. News* rankings over a period of six years (1995–2000) taking into account the changes in methodology for the business, education, engineering, law, and medical disciplines by correlating their 1995 and 2000 rankings. The correlations were strong, ranging from .72 for education to .92 for law; engineering was quite high at .88 (Clarke, 2001). In fact, examination of the data used in this study for the 2003, 2004, and 2005 *U.S. News* rankings for engineering colleges offering graduate programs revealed that the top 20 engineering colleges remained stable over this three-year period. This phenomenon is not limited to only the *U.S. News* college rankings. The rankings of America's leading universities have not changed much over time; they have remained relatively stable from Cattell's rankings in 1910 to the National Academy of Science's rankings in 1982 (Webster, 1992b). It is perhaps not surprising that top ranking institutions remain stable over time. In the *U.S. News* ranking algorithm, academic reputation is either the only indicator or heavily weighted among the indicators. There is also the media's obsession with perpetuating the image of the Ivy League schools as the shining examples against which the public should compare all other universities (Moll & Wright, 1998).

Ridley et al. (2001) examined the transition between tiers using the *U.S. News* ranking data for six years (1996–2001) for 162 four-year liberal arts colleges and found that most of the institutions remained in the original tier in a year-to-year comparison; only 14% of the institutions showed tier movement over the six-year period. Similar results were obtained in this study using the 2004 and 2005 *U.S. News* rankings for graduate engineering programs; only 11.1% of the 81 ranked engineering colleges showed tier movement from 2004 to 2005 (Table 19 in Chapter 4). Nevertheless, Ridley et al. (2001) recommended that institutions trying to improve their standing in the rankings should not be discouraged from making improvements that would be beneficial to the institution. Moreover, it was important for institutions to not only know their tier ranking but also their position within the tier to determine the distance to the next border, and especially around tier borders that tended to show the most volatility (Ridley et al., 2001).

Academic Reputation

Attempts to rank the academic quality of universities have roused great ire in the past and will continue to do so in the future (Webster, 1986, 1992b). College rankings have been around since the late nineteenth century (Stuart, 1995). Reputational rankings for graduate programs in American colleges and universities have been used in the ranking methodology of studies conducted by the "inventor of reputational rankings," Raymond Hughes (Stuart, 1995), to rank graduate school departments in 1925 (Webster, 1986); Kenneth Roose and Charles Andersen's 1970 ratings of graduate programs (Roose & Andersen, 1970; Webster, 1986); and the National Academy of Science's 1982 assessment of research-doctorate programs (National Research Council, 2004), considered the "Rolls Royce" of academic quality rankings (Webster, 1986). Although academic quality rankings have been around for almost 100 years, college and university administrators dislike the immensely popular *U.S. News* college rankings (when rankings are down) because they affect the size and quality of their applicant pool (Webster, 1992a). Academic reputation rater scores are heavily weighted within the *U.S. News* ranking algorithm (*U.S. News,* 2003; Monks & Ehrenberg, 1999, 2003) because graduates from renowned colleges and universities are more likely to have an edge in the job market for high-paying jobs or admission into top graduate schools over graduates from lesser-known institutions (Morse & Gilbert, 1995; Ehrenberg, 2003).

A study by McGuire (1995) on weights for the indicators in the 1992 *U.S. News* ranking methodology used to rank 140 national liberal arts colleges revealed that when he changed the weights based on the average weights recommended by a panel of experts, 88% of the colleges changed ranks, with some colleges moving by as much as 24 places with an average shift of 5 places, with the most movement occurring in the second and third quartiles. Academic reputation was more heavily weighted by *U.S. News* than by the panel of experts (McGuire, 1995). McGuire (1995) concluded that while there were no optimal weights for the ranking methodology (even his panel of experts disagreed on a consistent set of weights), "empirically derived weights are better than arbitrary ones" (p. 53). *U.S. News*, however, does solicit input from the higher education community; *U.S. News* staff meets with an average of 2.5 senior college officials per week in their Washington D.C., office and attend higher education conferences (Morse & Gilbert, 1995).

Despite criticisms of academic quality rankings, they are needed to provide comparative information on the quality of education in American universities (Webster, 1986). While the senior college officials who respond to the annual *U.S. News* surveys to rate academic quality are not familiar with every college on the survey, their opinion represents a collective consensus (Morse & Gilbert, 1995), albeit a subjective one. Universities and colleges do not presume to publish academic reputational rankings, perhaps because there is no consensus on how to objectively define quality in higher education vis-à-vis research quality, institutional quality, or teaching quality (Drew & Karpf, 1981) — nor how to measure academic quality (Stuart, 1995). Consulting college rankings and guidebooks can help parents and students make a more informed decision, since "The benefits of higher education are intangible, and not immediate" (Hossler, 1998, p. 165) and "Families love them because of the simplicity of the single vertical column that says college number 5 is better than number 6" (Moll & Wright, 1998, p. 158). In short, a good reputation increases the credibility of the organization in the eyes of the public (Fombrun, 1996).

It was in the 1970s and 1980s, when intense public scrutiny and declining college enrollment forced colleges to adopt business marketing and advertising strategies, that college ranking and guidebook publications became a booming market available to the general public (Hunter, 1995). The reputation of schools and colleges experienced a severe decline in the eyes of the public when the influential and highly critical publication, *Nation at Risk,* was released by the National Commission on Excellence in Education in 1983 (Clarke, 2001). The report warned of the impending jeopardy of America's position as a superpower in the world global marketplace if the American educational system was not reformed to be competitive with other leading nations (Bracey, 2003). Since then, schools and colleges have been trying to reverse public opinion. Savvy colleges and universities of today employ sophisticated marketing and business tools, ranging from videos, telemarketing, targeted mailings, college ranking publications, college guidebooks, to television, radio, and billboard advertisements, to build a reputation and attract students (Hossler & Foley, 1995; McGuire, 1995). Colleges and universities can "make themselves hot with some savvy self-promotion" (Shea, 2004b, p. 57), and a new wave of self-promotion can be seen on reality TV shows about student life on campus (Shea, 2004b).

According to Fombrun (1996), "The proliferation of such subjective rankings as 'best managed,' 'most innovative,' and 'most admired' attests to the growing popularity of reputation as a tool for assessing companies" (p. 6). In fact, Americans are fascinated by rankings: one can find rankings of sports teams, tall buildings, hospitals, and cities with the best business climate, among other things (Hossler, 2000). And the media continues to perpetuate the image of the elite and highly visible Ivy League colleges, but for the "equally pricey but not-so-well-knowns" (Moll & Wright, 1998, p. 159), one rarely reads about them in the media. According to Moll and Wright (1998):

> . . .these are the colleges that, lacking visibility and therefore positive image, are scrambling for students, worrying about net tuition revenue, and offering fire-sale prices to fill the beds. . .(p. 159) . . .One theory is that the media know that Americans are tremendously status conscious, with advertising and brand name a major aspect of capitalism. (p. 157)

Summary

This review of literature provided a brief history of the origins of college rankings; a description of the role of higher education in the global marketplace, the cost of higher education, and the importance of college rankings as an effective marketing tool; and the use of college rankings as an assessment tool in college choice and funding decisions. The chapter concluded with an overview of the *U.S. News* college rankings and of the impact of academic reputation as an assessment tool for educational quality.

CHAPTER 3

METHODOLOGY

Introduction

The *U.S. News* graduate engineering methodology uses *z* scores to determine overall rankings (see Table 1 for methodology). The *U.S. News* engineering methodology was replicated by calculating *z* scores for 87 institutions published in the 2005 *U.S. News* rankings across 11 indicators, by copying the data into Excel and setting up Excel worksheets to calculate the *z* scores. Examination of the 2003, 2004, and 2005 *U.S. News* data revealed that very few institutions had moved up or down in the rankings from year to year. In fact, as was noted previously in Chapter 2, *Review of Literature* section, the top 20 institutions consistently remained in the top tier. Two models were created: (a) a *U.S. News* replicate model to predict discrete changes in position order in the rankings, and (b) a tier model (pBAP) using the *U.S. News* ranking data to predict tier movement in the rankings. The words *case* and *institution* were used interchangeably throughout the *Methodology* (Chapter 3) and *Results* (Chapter 4) chapters and referred to Engineering colleges offering graduate programs. Likewise, the words *model, formula, methodology,* and *algorithm* were used interchangeably throughout the *Methodology* and *Results* chapters and referred to the *U.S. News* ranking model for engineering colleges offering graduate programs.

Ridley et al. (2001) recommended that it was important for institutions to not only know their tier ranking but also their position within the tier to determine the distance to next border, and especially around tier borders that tended to show the most volatility. Observation of the movement of institutions from year to year (2003–2004, and 2004–2005) showed that institutions were likely to move up or down in the rankings in discrete changes. A user-friendly model to predict movement over time from tier to tier was needed for colleges to determine their standing in relation to the top 20 and peer institutions. The Predictive Model for Benchmarking Academic programs (pBAP) was developed to fill this need. If deans of Colleges of Engineering had a goal to be in the top 20 colleges in the *U.S. News* rankings within a specified time period, for example, the pBAP model would provide a tool to predict movement by tier instead of discrete movement in the rankings. Classification discriminant analysis (DA) was deemed the appropriate statistical test to create the pBAP model. DA is the appropriate procedure when the dependent variables are categorical and the independent variables are metric; there are three or more classification groups; and there is correlation among the independent variables (Hair et al., 1998). Hair et al. (1998) also suggested that a minimum of 20 cases per group should be used, and an equal number of cases should be maintained within each group to avoid overclassification into the larger groups. A replicate model of the *U.S. News* methodology was also developed to predict discrete changes in position order in the rankings.

Table 1

2005 *U.S. News* Methodology for Engineering Colleges Offering Graduate Programs

Quality assessment (weighted by .40)	Peer score of engineering school deans and deans of graduate studies (weighted by .25) Recruiter score of corporate recruiters and company contacts from previously ranked programs (weighted by .15) Peer and recruiter scores are tabulated from survey respondents who are asked to rate program quality from marginal (1) to outstanding (5).
Student selectivity (weighted by .10):	Average quantitative GRE scores of master's and doctoral students (weighted by .45) Average analytical GRE scores of master's and doctoral students (weighted by .45) Acceptance rate of master's and doctoral students (weighted by .10)
Faculty resources (weighted by .25)	Ratio of full-time doctoral students to full-time faculty (weighted by .30) Ratio of full-time master's students to full-time faculty (weighted by .15) Proportion of full-time faculty in the National Academy of Engineering (weighted by .30) Number of doctoral degrees granted in last school year (weighted by .25)
Research activity (weighted by .25)	Total externally funded engineering research expenditures (weighted by .60) Research dollars per faculty member engaged in research (weighted at .40) Expenditures refer to separately funded research, public and private, conducted by the school, averaged over the last two fiscal years.

Overall rank: Data were standardized about their means, and standardized scores were weighted, totaled, and rescaled so that the top-scoring school received 100; others received their percentage of the top score. (Through personal communication, June 2004, with Samuel Flanigan, Deputy Director of Data Research, *U.S. News*, clarification was obtained on rescaling the *z* scores. An additional step was needed to more accurately replicate the rankings. To replicate overall rank percentages, *z* score values were rescaled to range from 0 to 100.)

Source: www.usnews.com

Statement of the Problem and Purpose of the Study

Improving national ranking is an increasingly important issue for university administrators (Arnone, 2003; Kleiner, 2004; Stecklow, 1995; Karl, 1999; Mallette, 1995; Dichev, 2001; Hunter, 1995; Thompson, 2000; Hossler, 2000; Ridley et al., 2001). While research has been conducted on performance measures in higher education, research designs have lacked a predictive quality. According to Doerfel and Ruben (2002):

> As with business, higher education indicators have tended to be primarily historical, limited in predictive power, often incapable of alerting institutions to changes in time to respond, and lacking adequate attention to important but difficult-to-quantify dimensions. Ironically, the emphasis on easy-to-quantify, limited measures has, in a manner of speaking, "come to haunt" in the form of popularized college rating systems with which educators generally are frustrated and critical. But these are used consistently as the measures against which universities are evaluated by their constituents. (p. 20)

The purpose of this study was to develop a predictive model for benchmarking academic programs (pBAP) for engineering colleges offering graduate programs based on the *U.S. News* college ranking methodology. This study was not concerned with the validity of the *U.S. News* methodology, but rather with the development of a predictive model within the existing *U.S. News* methodology. The pBAP model can be used to predict what changes in values would cause a college to move up or down a tier in ranking over time. A replicate study of the *U.S. News* methodology was also developed in this study to predict discrete changes in a college's ranking from year to year.

Data Sets of Interest

The predictive model developed in this study was based on the 2005 *U.S. News* data for engineering colleges offering graduate programs. This data set contained published rankings for 87 engineering colleges.. The 2005 *U.S. News* data were used to develop the pBAP model. To test the predictive accuracy of the pBAP model, the 2005 *U.S. News* data were entered into the pBAP variate developed using the 2004 *U.S. News* data, which contained data on 84 engineering colleges. Although a data set was available for the 2003 *U.S. News* rankings, it contained rankings for only 50 engineering colleges, which was an insufficient number of cases to create the four-tier pBAP model. Thus, these data were not used in the model. Instead, the 2003 *U.S. News* data set was used to examine the stability of the 20 top-ranked engineering colleges over a three-year period (2003–2005) and not in the predictive model. The data were collected by purchasing online premium access to the *U.S. News* Web site and were stored each year in an Excel file by the researcher. The *U.S. News* variables and methodology are listed in Table 1.

Assumptions of Classification Discriminant Analysis

Hair et al. (1998) cited two key assumptions for classification discriminant analysis (DA): (a) multivariate normality of the independent (discriminating) variables, and (b) equality of the covariance matrices for each group. DA is sensitive to departures from multivariate normality and equal covariances for the groups, but can be ignored and the results interpreted if certain conditions are met (Hair et al., 1998; Statsoft Inc., 2004). Violation of these assumptions is considered minimal if groups are approximately equal in size, and "if the largest group size divided by the smallest group size is less than 1.5" (Hair et al., 1998, p. 348). Other steps should be taken to improve the homogeneity violation. Independent variables not meeting normality assumptions can be transformed by computing the log 10, square root, inverse, or square root values of the variables

(Schwab, 2003). If the transformed variables meet the normality assumption, then they can be substituted in the data set to determine if the Box's *M* statistic has improved as well as the cross-validated classification accuracy. Outlier variables can also be examined and kept or removed if appropriate (Hair et al., 1998).

Research Design and Data Analysis

Steps in Replicating the U.S. News Model in Excel

First, the 2005 *U.S. News* data for the 87 ranked institutions were copied into a worksheet in an Excel workbook. Four of the 87 institutions did not have data for one or two variables denoted by "n/a" in the *U.S. News* ranking table; the "n/a" value was replaced by a zero value in the study. Then *U.S. News* row and column labels were copied into the second worksheet, which was set up with formulas to calculate *z* scores for each indicator within each case. Two extra columns were added to the second worksheet, *replicated overall z score* and *rescaled replicated overall score*, and two additional rows, *mean* and *standard deviation*. The mean and standard deviation for all cases for each indicator were calculated using the Excel functions, AVERAGE and STDEV, which were used to calculate the *z* score for each indicator within each case. Each indicator *z* score was multiplied by the assigned weight in the *U.S. News* methodology described in Table 1 and summed to produce the overall *z* score for each case. Equation 2 is the *U.S. News* algorithm.

$$.25\ peer + .15\ recrut + .10\ (.45\ quant + .45\ analyt + .10\ accept) + .25\ (.30\ phdfac + .15\ msfac + .30\ nae + .25\ phdno) + .25\ (.60\ rschex + .40\ rscfac) \quad (2)$$

where

peer = peer assessment score

recrut = recruiter assessment score

quant = average quantitative GRE score master's & doctoral students

analyt = average analytical GRE score master's & doctoral students

accept = acceptance (selectivity) rate master's & doctoral students

phdfac = ratio of full-time doctoral students to full-time faculty

msfac = ratio of full-time master's students to full-time faculty

nae = percentage of full-time faculty in the National Academy of Engineering

phdno = number of doctoral degrees granted for the last academic year

rschex = total externally funded engineering research expenditures

rscfac = research dollars per research-active faculty

The overall *z* score for each institution was rescaled by assigning 100% and 0% to the top (maximum) and bottom (minimum) overall *z* score respectively, using the formula in Equation 3. Institutions in between received their percentage of the top (maximum) overall *z* score. The resulting percentages were sorted in descending order to determine rank position order for each institution.

$$\frac{\textit{overall z score} - \textit{minimum overall z score}}{\textit{maximum overall z score} + |\textit{minimum}|\ \textit{overall z score}} \quad (3)$$

In a third Excel worksheet, the position order of institutions in the replicate rankings was compared to their order in the original rankings to verify the accuracy of the replicate rankings. The differences in position order between the replicate and original rankings were calculated in Excel, and the frequency of the differences in position order changes were summarized in SPSS.

Steps in Predicting Discrete Changes in the U.S. News Model in Excel

Because data on the 2004 *U.S. News* rankings were available, it was possible to determine the predictive power of the replicate *U.S. News* model. The steps in the previous section, *Steps in Replicating the U.S. News Rankings in Excel,* were followed to replicate the 2004 *U.S. News* rankings. Then the variable values for each institution from the 2005 *U.S. News* rankings were entered into the 2004 *U.S. News* replicate ranking model following a stepwise input process. The values for the other institutions remained constant, but the mean and standard deviation values were recalculated for each indicator with each entry. The overall *z* score for each institution was rescaled for each entry by assigning 100% and 0% to the top (maximum) and bottom (minimum) overall *z* score respectively. Institutions in between received their percentage of the top (maximum) overall *z* score. The resulting percentages were sorted in descending order to determine rank position order for each institution. The predicted 2005 *U.S. News* position order was compared with the actual 2005 *U.S. News* position order to determine the predictive accuracy.

Steps in Creating an Interactive Web-Based Rank Calculator: U.S. News Model

A user-friendly, interactive rank calculator is being created by a computer science master's student in asp.NET programming language. Engineering college administrators will be able to change variable values of the *U.S. News* indicators and to determine what changes in variable values would cause the college to move up or down in their position order in the rankings. The Web site will be secured by a password access. Upon logging in, the college administrator will be able to select a *U.S. News*-ranked institution from a pop-up box to edit variable values for that institution. Entering changed values for the selected variables and choosing the *Calculate* button will return the predicted rank position order. College administrators will also be able to save the session with the changed values. The variable values revert to the original values when the user starts a new session or logs out.

The programmer was given the 2005 *U.S. News* data for the published ranking of the top 87 engineering colleges to upload with the following steps to calculate and rescale the overall *z* score for the database. This model is the replicate of the *U.S. News* model.

1. Calculate the mean and standard deviation for each indicator.
2. Calculate the *z* scores for each indicator value (Equation 1 in Chapter 1).
3. Apply the *U.S. News* algorithm (Equation 2), and total the overall *z* score.
4. Rescale the overall *z* score and re-sort in descending order.
5. Display the change in position order.
6. Relate the mean, standard deviation, overall *z* scores, rescaled overall score, and position order for changes in indicator values when editing for prediction.
7. Reset to the original scores upon log out.
8. Save changes to indicator values in a folder.

Steps in Creating the Four-Tier Model (pBAP) Using the U.S. News Data

The 2005 *U.S. News* data were copied into SPSS. The 87 institutions were coded into one of four equal tiers with a minimum of 20 institutions per tier. Four of the 87 institutions did not have data for 1 or 2 (of 11) variables denoted by "n/a" in the *U.S. News* ranking table; the "n/a" value was replaced by a zero value in the study. The classification discriminant analysis (DA) test was executed in SPSS. Interpretation of the DA output was conducted as shown in the following steps:

1. The Cross-Validated Classification Results table was examined to determine how many cases were correctly classified into the tiers.
2. The Box's *M* statistic was examined to determine homogeneity of the covariance matrices.
3. The canonical correlation was examined to determine how much each function explained the variance of the indicators.
4. The unstandardized coefficients from the canonical discriminant table were used to create weights for the variables in the discriminant variate, which when summed produced the discriminant score for each case. The standardized coefficients were examined to explain the unique contribution of each independent variable to the discriminant function.
5. The cutoff score between tiers was determined by using the values from the group centroid table.
6. The structure matrix coefficients were examined to determine which indicators contributed most to the discriminant function.
7. The proportional chance criterion was calculated to determine the extent to which the cross-validated classification accuracy rate in the model exceeded the chance occurrence classification rate.
8. The Press' *Q* Statistic was calculated to determine if the cross-validated classification rate of the model was statistically significant relative to the chance occurrence classification rate.

9. If the Box's *M* statistic were statistically significant, the independent variables should be examined to determine which ones did not meet normality assumptions.
10. The independent variables values that did not meet normality assumptions (outliers) should be transformed by various mathematical computations (log 10, square root, inverse, and square) and substituted into the dataset using the IDF.CHISQ[p,df] function in SPSS.
11. Cases that were outliers should be identified by comparing the Mahalanobis D^2 value of each case in the Casewise Statistics table to a critical value for a probability of .01 with the degrees of freedom equal to the number of independent variables used to compute the Mahalanobis D^2.
12. Data sets should be run containing various combinations of the log transformations of independent variables not meeting normality assumptions and removal of outlier cases.

Steps in Predicting Tier Changes in the Four-Tier Model (pBAP)

Because data on the 2004 *U.S. News* rankings were available, it was possible to determine the predictive power of the pBAP model that was developed using the 2005 *U.S. News* data. An Excel worksheet linking the 2005 *U.S. News* data to the 2004 *U.S. News* pBAP variate was created. Then each 2005 *U.S. News* unstandardized variable value was multiplied by the 2004 weights from the canonical discriminant table for function one from the SPSS Classification Discriminant Analysis output. The cross product of each unstandardized indicator value and its respective weight was summed to produce the overall discriminant score for each case. Cutoff scores for the tiers were calculated using the group centroid values for function one from the SPSS DA output. Each case was assigned to a tier according to its discriminant score. The predicted tier assignment was compared with the original tier assignment for the 2005 *U.S. News* rank position order to determine the predictive accuracy of the model. Cases that were near tier borders were noted, since the line that is drawn between tiers is an arbitrary line (but an arbitrary line based on the recommended sample size for the DA test).

Limitations and Delimitations

Three limitations derive from the current model to predict the following year's rankings: (1) assumption that data for all other institutions other than the selected institution remain constant when predicting future *U.S. News* rankings; (2) potential changes to the *U.S. News* ranking algorithm; and (3) missing data from *U.S. News* non-ranked (not published) institutions. According to Robert Morse, Director of Data Research, *U.S. News*, changes are made to its ranking algorithms periodically in response to suggestions for improvement from the higher education community (Morse & Gilbert, 1995). *U.S. News* (2003) cautions that rankings are not necessarily comparable from year to year; therefore, the data on the colleges and not necessarily the rankings should be compared for year-to-year comparisons. A study by Clarke (2001) compared the 1995 and 2000 *U.S. News* graduate rankings for engineering and found a .88 correlation, taking into account changes to the methodology, thereby demonstrating that changes to the engineering methodology did not have a significant impact on the rankings. Based on the findings of Clarke's (2001) study, future changes by *U.S. News* to the engineering ranking methodology should have minimal impact on the rankings from the previous year. As a precaution, however, the discriminant analysis procedure should be rerun with the current data each year to determine the new weights for the pBAP model for the following year's predictions.

Summary

The predictive model (pBAP) developed in this study provides insight on the *U.S. News* ranking methodology for engineering colleges for the institutional researcher. However, for the model to be very useful to the busy university administrator, a user-friendly, interactive version of the model was needed. An interactive rank calculator was created by a computer science master's student in asp.NET programming language for the pBAP model. The programmer was given the 2005 *U.S. News* data, the discriminant variate based on the 2005 *U.S. News* data, the predetermined tiers, and the cutoff scores between tiers. Engineering college administrators will be able to change variable values of the *U.S. News* indicators and to determine what changes in variable values would cause the college to move up or down a tier in the rankings. The Web site is secured by a password access. Upon logging in, the college administrator selects a *U.S. News*-ranked institution from a pop-up box to edit. Entering changed values for the selected variables and choosing the *Calculate* button will return the predicted tier membership. College administrators will eventually also be able to save the session with the changed values. The variable values revert to the original values when the user starts a new session or logs out. The programmer is also currently working on a rank calculator to predict discrete changes in an institution's ranking based on the replicate *U.S. News* ranking methodology.

CHAPTER 4

ANALYSIS OF DATA

Introduction

The results of the replication study of the *U.S. News* ranking methodology for engineering colleges offering graduate programs and the discriminant analysis test are presented in this chapter. The success of the *U.S. News* methodology replication study and its ability to predict the following year's rankings are discussed. The steps used to create the pBAP equation (Equation 6) and the ability of the pBAP model to predict the following year's tier placement in the rankings are demonstrated. The three research questions are addressed. The three research questions were:

1. Which *U.S. News* indicators contributed the most to the prediction model to discriminate between rankings or tiers?
2. What phenomena influenced *U.S. News* peer score assessment of academic reputation?
3. What was the predictive accuracy of the replicate *U.S. News* model and the pBAP model?

Results

U.S. News Model Replication Study Results

Because data on only 87 of the 167 institutions that participated in the 2005 *U.S. News* survey were available to calculate the means and the standard deviations necessary to calculate the z scores, it was not possible to replicate the *U.S. News* rankings with 100% accuracy. Position order refers to an institution's actual sequential order in the rankings (ignoring ranking ties), and rank order refers to an institution's actual rank number (including ranking ties). For the 2005 *U.S. News* data, there were 87 institutions but 81 rankings when accounting for ranking ties. For the 2004 *U.S. News* data, there were 84 institutions but 83 rankings when accounting for ranking ties. *U.S. News* rankings list institutions that tie in ranking in alphabetical order within the ranking tie. About half of the 2005 *U.S. News* rankings had colleges that tied with two, three, or four other colleges within a ranking number. Thus, a 0, ± 1, ± 2, ± 3, or ± 4 range of change in position order was deemed appropriate in the replication study. Seventy-one of the 87 institutions (or 81.6%) had a 0, ± 1, ± 2, ± 3, or ± 4 change in position order. Overall, the change in position order ranged from -35 to 14.

U.S. News data for 84 institutions were also available for 2004 from a data file saved by the researcher from the previous year, with the exception of one indicator, the ratio of full-time master's students to full-time faculty. *U.S. News* does not publish data for this indicator in the ranking table; data for this indicator can only be obtained by manually searching for these data on the *U.S. News* site institution by institution (a tedious process) and computing the ratio. However, the researcher collected values for all indicators for the 2005 *U.S. News* data set, including the full-time master's to full-time faculty ratio indicator. This was possible because the 2005 data are currently published on the *U.S. News* Web site, whereas the 2004 data are not. Replication of the *U.S.* News engineering methodology using the 2004 survey data for 84 ranked engineering colleges, yielded a 0, ± 1, ± 2, ± 3, or ± 4 change in position order for 65 out of 84 institutions (or 77.4%). Overall, the change in position order ranged from -14 to 19.

Examination of the results of the two replication studies (2004 and 2005 *U.S. News* data) revealed similar consistency (77.4% and 81.6%, respectively) in the ability to duplicate the *U.S. News* engineering methodology, even with the missing full-time master's students to full-time faculty ratio indicator values in the 2004 data. This finding was consistent with the structure matrix output from the Classification Discriminant Analysis (DA) procedure run in SPSS using the 2005 *U.S. News* data. The full-time master's to full-time faculty ratio variable (MSFAC05) contributed the least to discriminant function one (Table 2), which will be the only function interpreted for the pBAP model because functions two and three were not statistically significant (Table 10). Likewise, the test of equality of group means table (Table 3) revealed that all variables were statistically significant except the full-time master's to full-time faculty ratio variable (MSFAC05), which one would consider dropping from the model.

Table 2

Structure Matrix SPSS Output from the DA Test Using 2005 *U.S. News* Data

	Function 1	Function 2	Function 3
PEER05	.660*	.061	-.012
RECRUT05	.592*	.021	.044
PHDFAC05	.372*	-.256	-.020
NAE05	.323*	.151	.089
RSCFAC05	.303*	.095	.036
ACCEPT05	-.140	.714*	.090
RSCHEX05	.382	.519*	-.197
PHDNO05	.350	.397*	-.032
MSFAC05	-.014	.374	-.481*
ANALYT05	.249	-.102	.280*
QUANT05	.240	-.165	.253*

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions. Variables ordered by absolute size of correlation within function.
*Largest absolute correlation between each variable and any discriminant function.

Table 3

Tests of Equality of Group Means Using 2005 *U.S. News* Data

Variable	Wilks' Lambda	F	df1	df2	Sig.
PEER05	.197	107.212	3	79	.000
RECRUT05	.234	86.399	3	79	.000
QUANT05	.643	14.629	3	79	.000
ANALYT05	.627	15.656	3	79	.000
ACCEPT05	.756	8.479	3	79	.000
PHDFAC05	.432	34.602	3	79	.000
NAE05	.504	25.902	3	79	.000
RSCHEX05	.410	37.930	3	79	.000
RSCFAC05	.537	22.659	3	79	.000
PHDNO05	.458	31.225	3	79	.000
MSFAC05	.928	2.045	3	79	.114

U.S. News Model Prediction Study Results

To test the prediction accuracy of the replicate *U.S. News* model, 2005 institution data were entered stepwise (one institution at a time) into the 2004 model while keeping the data for the other institutions constant. The resulting overall *z* scores were rescaled for each entry. Because the 2004 data did not contain values for the full-time master's to full-time faculty ratio, this indicator was excluded from both the 2004 and 2005 data sets. Data on only 82 of the 87 institutions were available, because five institutions ranked in 2005 were not ranked in 2004. The difference between the predicted position order and the actual position order for the 2005 *U.S. News* rank position order resulted in 70 out of 82 institutions (or 85.4%) correctly placed within 0, ± 1, ± 2, ± 3, or ± 4 range (Table 4).

One should keep in mind that the replication of the 2005 *U.S. News* rank position order was 81.6% successful. Examination of the difference in actual changes in rank number for institutions for the 2004 and 2005 *U.S. News* rankings revealed that 70 out of 83 (or 84.3%) had moved up or down 0, ± 1, ± 2, ± 3, or ± 4 places in the rankings (Table 5). Comparison of the predicted position order movement (85.4%) and the actual rank number movement (84.3%) using 0, ± 1, ± 2, ± 3, or ± 4 movement in position order of institutions confirmed 0, ± 1, ± 2, ± 3, or ± 4 change in position order as a valid means of representing movement in the ranking numbers and to account for ranking ties.

Table 4

Difference in Position Order Between Predicted vs. Actual 2005 *U.S. News* Model

	Position Difference	Frequency	Valid Percent
Valid	-24.00	1	1.2
	-14.00	1	1.2
	-10.00	1	1.2
	-9.00	1	1.2
	-8.00	1	1.2
	-6.00	2	2.4
	-4.00	2	2.4
	-3.00	5	6.1
	-2.00	5	6.1
	-1.00	10	12.2
	.00	21	25.6
	1.00	17	20.7
	2.00	4	4.9
	3.00	5	6.1
	4.00	1	1.2
	5.00	2	2.4
	8.00	1	1.2
	16.00	1	1.2
	32.00	1	1.2
Total	82		100.0

Table 5

Difference in Actual Rank Number Between 2004 and 2005 *U.S. News* Model

	Rank	Frequency	Valid Percent
Valid	-9	1	1.2
	-8	2	2.4
	-6	3	3.6
	-5	1	1.2
	-4	5	6.0
	-3	5	6.0
	-2	4	4.8
	-1	11	13.3
	0	14	16.9
	1	10	12.0
	2	8	9.6
	3	7	8.4
	4	6	7.2
	5	2	2.4
	6	2	2.4
	7	1	1.2
	8	1	1.2
Total	83		100.0

pBAP Model Creation Results

The 87 institutions ranked in the 2005 *U.S. News* were divided into four tiers with a minimum of 20 cases per tier, and the data were analyzed using the Classification Discriminant Analysis (DA) statistical test in SPSS. DA exhibits strong sensitivity to the ratio of the sample size to the number of independent variables (Hair et al., 1998). Studies recommend that a minimum ratio of valid cases to the independent variables should be 5:1 with a preferred ratio of 20:1 (Hair et al., 1998). In addition, the sample size in each group should be greater than the number of independent variables and preferably contain a minimum of 20 cases per group, with an equal number of cases within each group to avoid overclassification into the larger groups (Hair et al., 1998). Another reason for dividing the tiers into a minimum of 20 cases per tier is that the American consumer likes comparison ratings like the "top ten" and "best of" (Rogers & Rogers, 1997). Observation of other rankings in American culture shows that "top 20" and "top 25" designations are not uncommon for comparisons of longer lists. Tiers 1–3 had 21 cases per tier, and tier 4 had 20 cases. For the pBAP model using the 2005 *U.S. News* data in this study, the ratio of valid cases to independent variables was 7.5:1 (where the number of valid cases = 83, the number of independent variables = 11), which satisfied the minimum ratio of 5:1. Each tier contained a minimum of 20 cases (where Tier 1 = 21 cases, Tier 2 = 21 cases, Tier 3 = 21 cases, and Tier 4 = 20 cases), which exceeded the number of independent variables (11), and therefore satisfied the minimum number of cases recommended per group (20).

SPSS excluded 4 of the 87 institutions from the analysis because those institutions had 1 or 2 variable values published as "n/a." The "n/a" values were not replaced with a zero value in the pBAP data set as in the replication study for the *U.S. News* since there were enough cases with complete data to build the pBAP model. While correlation (multicollinearity) was strong between the 11 *U.S. News* independent variables, a characteristic that made the DA test suitable to build the pBAP model (Table 6), no independent variable unduly impacted the explanatory power of another independent variable. Multicollinearity is identified by examining tolerance values; if the tolerance value of an independent variable is less than 0.10, then multicollinearity is indicated (Schwab, 2003). If there were a problem with multicollinearity of the independent variables that were entered together, SPSS output of the DA test would include a table called “Variables Failing Tolerance Test.” SPSS did not create an output table of variables failing tolerance; therefore, there was no multicollinearity problem in this DA test. Table 6 shows the Pearson correlation coefficients between the variables.

The Box’s *M* test was statistically significant at p<.000. This result indicated that groups differed across their covariance matrices, a violation of DA; however, as noted previously in the *Assumptions of Classification Discriminant Analysis* section in Chapter 3, the Box’s *M* test is sensitive to deviations from multivariate normality and to unequal covariances for the groups and can be robust even when violated (Hair et al., 1998; Statsoft Inc., 2004). The cross-validated classification table revealed that 81.9% of the cases were correctly classified into the groups (Table 12) for this first (baseline) model. Several steps are recommended to minimize the impact of violation of multivariate normality.

Table 6

Pearson Correlation of 2005 *U.S. News* 11 Indicators

		Peer	Recrut	Quant	Analyt	Accept	PhDfac	NAE	Rschex	Rscfac	PhDno	MSfac
Peer	Pearson	1.000										
	Sig. (2-tailed)	.										
	N	87										
Recrut	Pearson	.963**	1.000									
	Sig. (2-tailed)	.000	.									
	N	87	87									
Quant	Pearson	.566**	.534**	1.000								
	Sig. (2-tailed)	.000	.000	.								
	N	85	85	85								
Analyt	Pearson	.609**	.619**	.687**	1.000							
	Sig. (2-tailed)	.000	.000	.000	.							
	N	84	84	84	84							
Accept	Pearson	-.339**	-.305**	-.509**	-.471**	1.000						
	Sig. (2-tailed)	.001	.004	.000	.000	.						
	N	86	86	84	83	86						
PhDfac	Pearson	.705**	.664**	.509**	.478**	-.490**	1.000					
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.					
	N	87	87	85	84	86	87					
NAE	Pearson	.677**	.694**	.498**	.538**	-.225*	.664**	1.000				
	Sig. (2-tailed)	.000	.000	.000	.000	.037	.000	.				
	N	87	87	85	84	86	87	87				
Rschex	Pearson	.763**	.710**	.364**	.309**	-.187	.526**	.412**	1.000			
	Sig. (2-tailed)	.000	.000	.001	.004	.085	.000	.000	.			
	N	87	87	85	84	86	87	87	87			
Rscfac	Pearson	.479**	.477**	.323**	.267*	-.209	.514**	.525**	.620**	1.000		
	Sig. (2-tailed)	.000	.000	.003	.014	.053	.000	.000	.000	.		
	N	87	87	85	84	86	87	87	87	87		
PhDno	Pearson	.832**	.772**	.357**	.321**	-.123	.556**	.473**	.890**	.433**	1.000	
	Sig. (2-tailed)	.000	.000	.001	.003	.259	.000	.000	.000	.000	.	
	N	87	87	85	84	86	87	87	87	87	87	
MSfac	Pearson	-.096	-.145	-.051	-.333*	.504*	-.076	-.033	.101	.163	.128	1.000
	Sig. (2-tailed)	.375	.181	.641	.002	.000	.484	.758	.352	.132	.239	.
	N	87	87	85	84	86	87	87	87	87	87	87

** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

First, the researcher should keep group sizes equal when possible to avoid overclassification of cases into the larger group. A rule of thumb is if the ratio is less than 1.5, then violation of this assumption is considered to be minimal (Hair et al., 1998). For the pBAP model, the largest group size (21) divided by the smallest group size (20) was 1.05; therefore, the pBAP model met this criterion. Second, the researcher should examine each independent variable to determine which one(s) did not meet the normality assumption. The 2005 *U.S. News* data set was run using a test for normality assumption script in SPSS. The skewness and kurtosis (peakedness) of the data curve for each independent variable were examined to determine which of the 11 variables violated the normality assumption (Table 7). If the skewness or kurtosis values were outside the range of -1.0 to 1.0 in the charts, then the criteria for normal distribution of the data would not be satisfied. Six of the 11 independent variables did not meet the normal distribution criterion: quant, accept, nae, rschex, phdno, and msfac as shown in Table 7.

Table 7

Normality Assumption Tests for 11 Variables Using 2005 *U.S. News* Data

Variable Name*	Skewness	Kurtosis	Between -1.0 to 1.0	Normal?
PEER05	0.767	-0.006	yes (both)	yes
RECRUT05	0.690	0.270	yes (both)	yes
QUANT05	-1.124	3.058	no (both)	no
ANALYT05	-0.433	0.024	yes (both)	yes
ACCEPT05	1.092	1.139	no (both)	no
PHDFAC05	0.111	-0.555	yes (both)	yes
NAE05	1.436	1.820	no (both)	no
RSCHEX05	1.744	2.933	no (both)	no
RSCFAC05	0.955	0.861	yes (both)	yes
PHDNO05	1.555	2.092	no (both)	no
MSFAC05	1.946	5.040	no (both)	no

*See Tables 1 (Chapter 3) and Table 15 for the long description of variable names.

Third, the researcher should transform the variables not meeting the normality assumption in an attempt to normalize the data. The 2005 *U.S. News* data set was run using a transformation script in SPSS for the six independent variables not meeting the normality assumption mentioned above. The transformations used were log 10, square root, inverse, and square. The skewness and kurtosis were checked for each transformation type for the six variables (Table 8). Two of the six variables (nae and quant) did not meet the normal distribution range criteria of -1.0 to 1.0; therefore, they should be interpreted with caution. Four of the six variables (accept, rschex, phdno, and msfac) met the range for normal distribution with the log 10 transformations.

The Classification Discriminant Analysis (DA) test was rerun in SPSS with the 2005 *U.S. News* data using the original values for seven variables and substituting the log 10 transformation values for four variables (accept, rschex, phdno, and msfac). The original values for the two variables, quant and nae, were used in the analysis since the transformations in Table 8 did not normalize the data. The cross-validated classification accuracy was 78.3%, which was less than the 81.9% accuracy rate in the first (baseline) model. The Box's *M* statistic remained statistically significant at $p<.0000$ with the substitutions, indicating violation of homogeneity. This second (transformed) model was therefore rejected in favor of the first (baseline) model.

Table 8

Normality Assumption Tests for Transformations for 6 Variables

Variable: QUANT05

Transformation	Skewness	Kurtosis	Between -1.0 to 1.0	Normal?
Log 10	-2.337	10.658	no (both)	no
Square Root	-1.05	1.415	no (both)	no
Inverse	8.595	76.888	no (both)	no
Square	-0.983	2.491	no (kurtosis)	no

Variable: NAE05

Transformation	Skewness	Kurtosis	Between -1.0 to 1.0	Normal?
Log 10	1.339	1.461	no (both)	no
Square Root	1.389	1.635	no (both)	no
Inverse	-1.245	1.138	no (both)	no
Square	2.958	9.982	no (both)	no

Variable: ACCEPT05

Transformation	Skewness	Kurtosis	Between -1.0 to 1.0	Normal?
Log 10	-0.082	-0.569	yes (both)	yes
Square Root	0.491	-0.137	yes (both)	yes
Inverse	1.127	0.855	no (skewness)	no
Square	2.293	6.096	no (both)	no

Variable: RSCHEX05

Transformation	Skewness	Kurtosis	Between -1.0 to 1.0	Normal?
Log 10	0.115	-0.263	yes (both)	yes
Square Root	0.995	0.458	yes (both)	yes
Inverse	2.832	11.954	no (both)	no
Square	3.192	11.632	no (both)	no

Variable: PHDNO05

Transformation	Skewness	Kurtosis	Between -1.0 to 1.0	Normal?
Log 10	-0.214	0.025	yes (both)	yes
Square Root	0.790	0.160	yes (both)	yes
Inverse	3.573	17.536	no (both)	no
Square	2.784	8.297	no (both	no

Variable: MSFAC05

Transformation	Skewness	Kurtosis	Between -1.0 to 1.0	Normal?
Log 10	-0.468	0.835	yes (both)	yes
Square Root	0.797	1.212	no (kurtosis)	no
Inverse	3.473	14.542	no (both)	no
Square	3.966	19.030	no (both)	no

Fourth, the researcher should check for outliers among the cases. The squared Mahalonobis distance to centroid values in the Casewise Statistics table in the Discriminant Analysis output in SPSS were examined for outliers by comparing the table values to the critical value. The critical value for the Mahalanobis D^2 was 24.72, which was computed using the SPSS function IDF.CHISQ[p,df] where p = 0.99 (for a probability level of 0.01) and df = 11 (the number of variables used to compute D^2). Examination of the Casewise Statistics table revealed no outliers for the original classification computation but seven outliers (Cases 1, 2, 3, 6, 7, 19, and 68) for the cross-validated computation. For the third model, the 2005 *U.S. News* data set was rerun in SPSS for the DA procedure using the original values of the 11 independent variables but removing the seven outlier cases. The total number of validated cases was thereby reduced from N = 83 to N = 76, and each of the four tiers contained 19 cases each. The cross-validated classification rate was 81.6%, which was less than the 81.9% in the first (baseline) model. The Box's *M* statistic remained statistically significant at $p<.000$, indicating violation of homogeneity. This third model (removal of outliers) was rejected in favor of the first (baseline) model.

For the fourth model, the 2005 *U.S. News* data set was rerun in SPSS for the DA procedure using the original values of seven independent variables, the transformed (log 10) values of four independent variables, and removal of six outlier cases (cases 1, 7, 19, 68, 81, and 83). (Six outliers were indicated using log transformations versus seven outliers using all original variable values in the above third model.) The total number of validated cases were thereby reduced from N = 83 to N = 77, and the first tier contained 20 cases and tiers two–four contained 19 cases each. The cross-validated classification rate was 77.9%, which was less than the 81.9% in the first (baseline) model. The Box's *M* statistic remained statistically significant at p<.000, indicating violation of homogeneity. This fourth model (log 10 transformation substitution values and removal of outliers) was rejected in favor of the first (baseline) model.

For each of the four models, the Box's *M* statistic remained statistically significant at p<.000, indicating consistent violation of homogeneity. A final (fifth) model was run in SPSS using the DA test and the original values of the 11 independent variables but checking the separate-group covariance option for classification versus the within-group covariance option. The original classification correctly classified 95.2% of the cases into the groups; however, no cross-validated classification results were produced. Original classification results tend to overclassify the cases into the groups. The Box's *M* statistic remained statistically significant at p<.000, indicating violation of homogeneity. This final (fifth) model was also rejected in favor of the first (baseline) model. Table 9 shows the classification rate of each of the models.

Table 9

Five Models Tested to Minimize Violation of Homogeneity Using 2005 *U.S. News* Data

Model	Cross-Validated Classification	Homogeneity Violation?
First (baseline)	81.9%	yes
Second (log transformation)	78.3%	yes
Third (outliers removed)	81.6%	yes
Fourth (log transform + outliers)	77.9%	yes
Fifth (separate-group covariance)	(95.2%)*	yes

*Original classification; no cross-validated classification output.

The pBAP equation was derived from the canonical discriminant function coefficients of function one from the DA procedure using the baseline model (Table 15). Since there were four groups or tiers, three functions were derived. The Wilks' Lambda (λ = .065) and chi square (χ^2 =203.249) tests showed that function one was statistically significant at p<.000 (Table 10). The other two functions were not statistically significant; therefore, they were excluded from the model. The squared canonical correlation $(.95)^2$ for function one explained 90.25% of the variance in group membership (Table 11). The four-tier pBAP model also correctly classified 81.9% of the cases in their respective tier (Table 12). In the absence of any general guidelines to determine the degree to which a model should exceed the classification prediction accuracy by chance occurrence alone (Hair et al., 1998), a 25% improvement over the rate of accuracy achievable by chance alone is recommended to characterize a discriminant model as useful (Hair et al., 1998; Schwab, 2003).

To determine the predictive power of the classification accuracy of 81.9% for the pBAP model relative to correct classification of cases by chance occurrence alone, the proportional by chance accuracy rate was calculated by squaring the prior probabilities values for each group in the "Prior Probabilities for Groups" table in the SPSS output, summing the results, and multiplying the total by 1.25 (Table 13). The proportional chance criterion randomly assigns cases proportionate to the number of cases in each group and is used when group sizes are unequal (Hair et al., 1998). The 81.9% classification rate showed a 162% improvement over a 25% increase in the proportional chance accuracy rate achievable by chance alone of 31.3%. Table 13 also compares each group's 25% increase over the proportional chance accuracy rate achievable by chance alone with each group's classification accuracy rate.

Table 10

Statistical Significance of Discriminant Functions 2005 *U.S. News* Data

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 3	0.065	203.249	33	0.000
2 through 3	0.676	29.12	20	0.085
3	0.857	11.472	9	0.245

Table 11

Accounting for Variable Variance of Each Discriminant Function 2005 *U.S. News* Data

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	9.353*	95.6	95.6	0.95
2	0.267*	2.7	98.3	0.459
3	0.166*	1.7	100	0.378

*First 3 canonical discriminant functions were used in the analysis.

Table 12

Cross-Validated Tier Classification Results 2005 *U.S. News* Data

Yr: 2005		Predicted Group Membership				
Type	*Tier*	*1*	*2*	*3*	*4*	*Totals*
Count	*1*	*18*	3	0	0	21
	2	1	*16*	4	0	21
	3	0	2	*16*	3	21
	4	0	0	2	*18*	20
Percentage	*1*	*85.7%*	14.3%	0.0%	0.0%	100%
	2	4.8%	*76.2%*	19.0%	0.0%	100%
	3	0.0%	9.5%	*76.2%*	14.3%	100%
	4	0.0%	0.0%	10.0%	*90.0%*	100%

81.9% of cross-validated grouped cases correctly classified. N=83; 4 cases with missing data were excluded. Proportional by chance accuracy rate (x 25% improvement) = 31.3% (see Table 10).

Table 13

25% Improvement Over Chance Classification vs. pBAP Model Classification Accuracy

GROUP05	Prior	Cases Used in Analysis Unweighted	Cases Used in Analysis Weighted	$Prior^2$	25% Over $Prior^2$	Classification*
1	.253	21	21.000	.064	*8.0%*	85.7%
2	.253	21	21.000	.064	*8.0%*	76.2%
3	.253	21	21.000	.064	*8.0%*	76.2%
4	.241	20	20.000	.058	*7.3%*	90.0%
Total	1.000	83	83.000	.250	*31.3%*	81.9%

*Cross-validated classification accuracy 81.9% (see Table 12).

To determine whether the 81.9% classification accuracy of the pBAP model was statistically significant relative to the 31.3% proportional chance accuracy rate with a 25% improvement over chance occurrence, the Press's *Q* statistic was computed. The equation to compute the *Q* statistic is (Hair et al., 1998):

$$Press' \ Q = \frac{[N - (nK)]^2}{N(K - 1)} \tag{4}$$

where

N = total sample size

n = number of observations correctly classified

K = number of groups

The Press' *Q* statistic is then compared to the critical value, defined as the chi-square (χ^2) value for K – 1 degrees of freedom (*df*) at the chosen confidence level (α) (Hair et al., 1998).

The value of the Press' *Q* statistic for the pBAP model (143.46) greatly exceeded the χ^2 critical value (11.35). Therefore, the discriminatory power of the pBAP model to correctly classify cases into the groups is significantly better than that of chance, thereby supporting cross-validation. The computation of the Press' *Q* statistic for the pBAP model is shown below.

$$\text{Press' } Q = \frac{[83 - (68 \times 4)]^2}{83(4-1)} = 143.46$$

where

N = 83

n = 68

K = 4

and critical value

$\chi^2_{(df=3,\ \alpha=.01)} = 11.35$

Press *Q* statistic 143.46 > critical value χ^2 11.35.

The critical value of χ^2 was retrieved from the table of critical values on the Web site of the National Institute of Standards (retrieved April 9, 2005).

The pBAP model was tested with an example of an institution ranked in the fourth tier in the 2005 *U.S. News* rankings to determine if the peer and recruiter assessment scores were increased, the institution would move up to the third tier. Peer and recruiter assessment score values for the test case institution were each increased by 0.4 in the data set. Peer and recruiter score values were chosen for change because the Pearson correlation between peer and recruiter scores was highly significant at .963 (Table 6). Thus, if peer score increased or decreased, it was likely that recruiter score would manifest the same magnitude and direction as the peer score and vice versa. The structure matrix table (Table 2) also showed that for discriminant function one, the peer and recruiter variables loaded the highest on and had the strongest correlation with the discriminant function. The following example shows how the pBAP model was created using the DA SPSS output tables, the 2005 *U.S. News* data, and the test case institution.

Example: Institution Ranked in the Fourth Tier

First, calculate the cutoff score between the tiers using the group centroid values for tiers 3 and 4 from the discriminant analysis (DA) SPSS output (Table 14).

Table 14

Group Centroids for the Discriminant Function 1

Tier	Function 1	Tier	Function 1	Tier	Function 1	Tier	Function 1
1	4.414	2	0.625	3	-1.520	4	-3.695

Equation 5 shows the calculation to determine the cutoff scores:

$$(centroid\ 1 + centroid\ 2) / 2 \qquad (5)$$

The cutoff score between tiers 3 and 4 is calculated as follows, using the group centroids in Table 14:

$$(-1.520 + -3.695) / 2 = (-5.215) / 2 = -2.6075$$

Next, create the equation to determine the school's discriminant score (H) using the canonical discriminant function coefficients (Table 15). If the discriminant score for the institution is greater than the cutoff score for the tier above, the institution moves to the next tier; if less than the cutoff score for the tier above, it remains in the current tier. The key for the short description variable names in Equation 6 that correspond with the

U.S. News graduate engineering variables is listed in Table 15. Equation 6 shows the resulting pBAP equation:

$$H = -21.932 + 2.95(peer) + 0.534(recrut) + 0.005(quant) + 0.006(analyt) - 0.627(accept) + 0.292(phdfac) + 0.051(msfac) - 0.793(nae) + 0.006(rschex) + 0.005(rscfac) - 0.008(phdno) \quad (6)$$

Table 15

Canonical Discriminant Function Coefficients (Unstandardized Coefficients)

Variable	Long Description Variable Name*	Coefficient
(Constant)		-21.932
PEER	peer assessment score	2.950
RECRUT	recruiter assessment score	0.534
QUANT	average quantitative GRE score	0.005
ANALYT	average analytic GRE score	0.006
ACCEPT	acceptance rate	-0.627
PHDFAC	ratio of full-time PhD students/faculty	0.292
MSFAC	ratio of full-time MS students/faculty	0.051
NAE	full-time faculty in the NAE	-0.793
RSCHEX	engineering research expenditures	0.006
RSCFAC	research dollars per faculty member	0.005
PHDNO	number of doctoral degrees granted	0.008

*See Table 1 in Chapter 3 for definitions of the 2005 *U.S. News* variables.

The discriminant score for the test case institution in the fourth tier was -3.67924. This score was calculated as follows:

$$-3.67924 = -21.932 + 2.95(2.2) + 0.534(2.6) + 0.005(705) + 0.006(615) - 0.627(.365) + 0.292(2.9) + 0.051(3) - 0.793(0) + 0.006(44.6) + 0.005(480) - 0.008(36)$$

Since -3.67924 is below the cutoff score of -2.6075 for the third and fourth tier, it belongs in the fourth tier.

For this example, a 0.4 increase in both the peer and recruiter scores, raising peer score from 2.2 to 2.6 and recruiter score from 2.6 to 3.0, is entered into the pBAP model. The new discriminant score resulting from the new values is -2.28564.

-2.28564 =
-21.932 + 2.95(*2.6*) + 0.534(*3.0*) + 0.005(705) + 0.006(615) – 0.627(.365)
+ 0.292(2.9) + 0.051(3) – 0.793(0) + 0.006(44.6) + 0.005(480) – 0.008(36**)**

This score is above the cutoff score of -2.6075 for the third and fourth tier; therefore, the institution moves to the third tier.

Steps in Calculating Discriminant Scores

A step-by-step calculation of the discriminant score follows.

Step 1: Identify the variables to change by using the shortcut method below to calculate by hand the H value of the original variables (anywhere from 1 to 11 variables; plug all new values into the equation if all 11 variables are changed). In this example, the values for the peer and recruiter score variables were changed.

Peer and Recruiter score (values are 2.2 and 2.6 respectively):

2.95(2.2) + 0.534(2.6) = 6.49 + 1.3884 = 7.8784

Step 2: Subtract the step 1 answer from the original H score from the tier ranking.

-3.67924 – 7.8784 = -11.55764*

*This value is the unchanged portion of the H value to be used in Step 4.

Step 3: Change the values of the variables chosen, and calculate the new piece of the H value. (Here, peer score was raised to 2.6 and recruiter score to 3.0, a 0.4 increase in both from the original scores.)

2.95(2.6) + 0.534(3.0) = 7.67 + 1.602 = 9.272

Step 4: Add the step 3 answer to the step 2 answer to find the new H value, and compare to the given cutoff between the original tier and the next highest tier.

$$9.272 + (-11.55764) = -2.28564$$

Result: -2.28564>-2.6075, which places the institution into the third tier from its current fourth tier position.

pBAP Model Prediction Results

Since data were available for the 2004 *U.S. News* rankings, the weights for the pBAP model could be calculated for 2004 and a prediction accuracy determined for the 2005 *U.S. News* rankings. The resulting 2004 pBAP model below (Equation 7) did not include the ratio of full-time master's students to full-time faculty. As previously noted, collecting data for this indicator involved a tedious process, and the researcher did not collect data for this indicator for the 2004 *U.S. News* rankings. However, the full-time master's to full-time faculty ratio variable contributed least to the discriminant function as shown in the structure matrix table (Table 2) using the 2005 *U.S. News* ranking data.

$$H = -19.225 + 3.766(peer) + -0.217(recrut) + 0.017(quant) - 0.010(analyt) - 1.564(accept) + 0.145(phdfac) - 1.250(nae) + 0.006(rschex) + 0.004(rscfac) - 0.004(phdno) \qquad (7)$$

The 2004 pBAP equation (Equation 7) was derived from the canonical discriminant function coefficients of function one from the classification discriminant analysis (DA) test using the 2004 *U.S. News* data. Since there were four tiers, three functions were derived. The Wilks' Lambda (λ = .065) and chi square (χ^2 =199.482) tests showed that function one was statistically significant at $p<.000$ (Table 16). Function two was also statistically significant at $p<.004$, Wilks' Lambda ($\lambda = .598$), and chi square ($\chi^2 = 37.541$). Function three was not statistically significant at $p<.699$. Both functions two and three were excluded from the model; although function two was statistically significant, its squared canonical correlation $(.596)^2$ explained only 35.5% of the variance in group membership, whereas function one explained 89.1% $(.944)^2$ of the variance (Table 17). In addition, the explained variance for function one in both the 2004 and 2005 models are comparable; function one explained 90.25% of the variance in group membership in the 2005 model (Table 11). The 2004 four-tier pBAP model also correctly classified 76.5% of the cases in their respective tier upon cross-validation (Table 18), compared with the 2005 model that correctly classified 81.9% of the cases upon cross-validation (Table 12). Similar to the 2005 test of equality of covariances for the groups, the Box's *M* was statistically significant at $p<.000$. Also similar to the 2005 model, the largest group size (21) divided by the smallest group size (20) was 1.05. Since 1.05 is less than the 1.5 rule-of-thumb ratio (Hair et al., 1998), violation of the assumption of equality of covariances for the groups was considered to be minimal, and it was deemed appropriate to proceed with the test. Since the results of the discriminant analysis test using the 2004 *U.S. News* data set demonstrated similar characteristics to the 2005 *U.S. News* data set, the researcher did not perform additional steps to check for normality of assumptions, to transform variable values, to compute the proportional chance criterion, nor to compute Press' *Q* statistic.

Table 16

Statistical Significance of Discriminant Functions 2004 *U.S. News* Data

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 3	.065	199.482	30	.000
2 through 3	.598	37.541	18	.004
3	.927	5.538	8	.699

Table 17

Accounting for Variable Variance of Each Discriminant Function 2004 *U.S. News* Data

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	8.192*	92.9	92.9	.944
2	.550*	6.2	99.1	.596
3	.079*	.9	100.0	.270

*First 3 canonical discriminant functions were used in the analysis.

Table 18

Cross-Validated Tier Classification Results 2004 *U.S. News* Data

Yr: 2004		Predicted Group Membership				
Type	*Tier*	*1*	*2*	*3*	*4*	*Totals*
Count	*1*	*15*	6	0	0	21
	2	0	*17*	3	0	20
	3	0	2	*15*	3	20
	4	0	0	5	*15*	20
Percentage	*1*	*71.4%*	28.6%	0.0%	0.0%	100%
	2	0.0%	*85.0%*	15.0%	0.0%	100%
	3	0.0%	10.0%	*75.0%*	15.0%	100%
	4	0.0%	0.0%	25.0%	*75.0%*	100%

76.5% of cross-validated grouped cases correctly classified. N=81; 3 cases with missing data were excluded.

The discriminant analysis function tends to overclassify cases into their groups, thereby resulting in a higher prediction accuracy than is valid (Hair et al., 1998). To remedy the overfitting of cases, SPSS cross-validates the classification of cases into their groups by omitting one case at a time from the original sample for subsequent classification of the omitted case (also known as jackknifing) (Hair et al., 1998). The result is a more rigorous estimate of the classification prediction accuracy. Another test that employs holdout and analysis samples to measure the internal consistency of the variables to determine if they are measuring the same underlying construct is Cronbach's Alpha reliability test (Brown, 2002). In the Cronbach's Alpha reliability test, the data are split into two halves, and all possible combinations are compared between the two halves and correlated to determine the strength of the relationship of items in both halves (National Society for Multiple Sclerosis, 2002). As the alpha value approaches 1, the stronger the indication of internal consistency, suggesting that the items are testing the same construct (Brown, 2002).

Applying Cronbach's Alpha test to the 2005 *U.S. News* data in SPSS for N = 83 cases and N = 11 independent variables resulted in α =.4958 and standardized α =.8397 (see Appendix A for the SPSS output). Applying the same test to the 2004 *U.S. News* data for N = 81 cases and N = 10 independent variables resulted in α =.4601 and standardized item α =.8319 (see Appendix B for the SPSS output). The reliability coefficients for the 2004 and 2005 *U.S. News* data were consistent. The regular alpha was less than the standard acceptable cutoff, α =.70, for the social sciences (Garson, retrieved July 9, 2004), while the standardized alpha was greater. The regular alpha is usually less than or equal to the standardized item alpha, and both the regular alpha and standardized item alpha are widely accepted (Garson, retrieved July 9, 2004). Examination of the coefficients of the regular and standardized alpha for the 2004 and 2005 *U.S. News* data revealed that the strength of the independent variables testing the same construct (quality of education) ranged from moderate (regular alpha) to strong (standardized alpha).

Table 19 shows the ability to use the current model to predict tier movement in the future, using the 2004 pBAP model and the 2004 *U.S. News* replicate model to predict institutions that demonstrate tier movement in the 2005 *U.S. News* rankings. Only a few institutions actually demonstrated tier movement in the following year; most of the institutions remained within the same tier. Tier movement from one year to the next was quite small: for the predictive models, 11.1% for the pBAP model and 13.6% for the replicate *U.S. News* model, compared with 12.3% actual tier movement (Table 19). This finding is consistent with a study by Ridley et al. (2001) in which they examined the transition between tiers using the *U.S. News* ranking data for six years (1996–2001) for 162 four-year liberal arts colleges and found that most of the institutions remained in the original tier in a year-to-year comparison; only 14% of the institutions showed tier movement over the six-year period. Nevertheless, Ridley et al. (2001) recommended that institutions trying to improve their standing in the rankings should not be discouraged from making improvements that would be beneficial to the institution. Moreover, it was important for institutions to not only know their tier ranking but also their position within the tier to determine the distance to the next border, especially around tier borders that tended to show the most volatility (Ridley et al., 2001). The pBAP model predicted 88.9% of the institutions would remain in the same ranking tier in the following year and the *U.S. News* replicate model 86.4%, compared with 87.7% of the institutions that actually remained in the same ranking tier (Table 19).

Observations by Peter Cary, *U.S. News* Special Projects Editor; Brian Duffy, *U.S. News* Executive Editor (1999); Dichev (2001); and Clarke (2001) further support this finding that the rankings of institutions in the top tier tend to remain stable over time. In fact, examination of the data used in this study for the 2003, 2004, and 2005 *U.S. News* rankings for engineering colleges offering graduate programs revealed that the top 20 engineering colleges remained stable over this three-year period. It is perhaps not surprising that top ranking institutions remain stable over time. In the *U.S. News* ranking algorithm, academic reputation is either the only indicator or heavily weighted among the indicators. There is also the media's obsession with perpetuating the image of the Ivy League schools as the shining examples against which the public should compare all other universities (Moll & Wright, 1998).

The pBAP model and the *U.S. News* replicate model were able to predict tier movement with approximately equal accuracy. However, the pBAP model is easier to calculate and use for tier prediction since the overall discriminant score for each case only involves substituting variable values into the linear equation, whereas the *U.S. News* model also involves recalculating the mean and standard deviation of each independent variable.

Table 19

Predicting Tier Movement from the 2004 *U.S. News* to the 2005 *U.S. News* Rankings

Movement	pBAP Predict Cases	*U.S. News* Predict Cases	Actual Cases
Remained in same tier	72 (88.9%)	70 (86.4%)	71 (87.7%)
Tier jump	9 (11.1%)	11 (13.6%)	10 (12.3%)
Total	81 (100%)	81 (100%)	81 (100%)

Tier Jump Characteristics			
Movement	pBAP Predict Cases	*U.S. News* Predict Cases	Actual Cases
Border tier jump*	3	3	4
Tier jump (not border)	6	8	6
Total	9**	11**	10**
Tier jump (exclude border)	6 (60%)	8 (73%)	

* Institutions that moved up or down 1 or 2 positions from the tier border from 2004 to 2005.
**Total of 15 unique institutions: 5 exclusive to the pBAP and *U.S. News* Predict Cases columns, 10 appearing across the pBAP and *U.S. News* Predict Cases and Actual Cases columns.
Note: The number of Predict Cases is based on ranking numbers (including ties) vs. rank position order in Tables 4 and 5. For the published 2005 *U.S. News* rankings, institutions were ranked through 81.

Summary

The answers to the three research questions are summarized in this section. The first research question asked which *U.S. News* variables contributed most to the prediction model variate to discriminate between discrete rankings or tiers. Comparison of the absolute value of the variable coefficients in the structure matrix tables from the discriminant analysis test for the 2004 and 2005 *U.S. News* data showed that the top three influential indicators of the discriminant function one consistently were: peer assessment score, recruiter assessment score, and research expenditures (Table 20). The structure matrix coefficients determine the contribution of each independent variable as influenced by the other independent variables (Statsoft Inc., 2004). Larger coefficients exert a stronger influence in explaining the variance in the function.

Table 20

Comparison of Structure Matrix Coefficients for the 2004 and 2005 *U.S. News* Data

2005 Indicator	2005 Function 1	2005 Contribution Rank Order	2004 Indicator	2004 Function 1	2004 Contribution Rank Order
PEER05	.660	1	PEER04	.734	1
RECRUT05	.592	2	RECRUT04	.512	2
PHDFAC05	.372	4	PHDFAC04	.364	4
NAE05	.323	6	NAE04	.333	5
RSCFAC05	.303	7	RSCFAC04	.251	6
ACCEPT05	-.140	10	ACCEPT04	-.251	6
RSCHEX05	.382	3	RSCHEX04	.433	3
PHDNO05	.350	5	PHDNO04	.364	4
MSFAC05	-.014	11	MSFAC04	not available	not available
ANALYT05	.249	8	ANALYT04	.200	8
QUANT05	.240	9	QUANT04	.232	7

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions.

Comparison of the absolute value of the standardized coefficients in the canonical discriminant function table (Table 21) from the discriminant analysis test for the 2004 and 2005 *U.S. News* data sets showed that the top two influential indicators of the discriminant function consistently were peer assessment score and research expenditures per faculty. Unlike the coefficients of the structure matrix that determine the contribution of each independent variable as influenced by the other independent variables with the discriminant function, the standardized coefficients of the canonical discriminant function explain the unique contribution of each independent variable to the discrimination between groups, in this case, ranking tiers (Hair et al., 1998). In practice, however, structure coefficients are more reliable in determining the influence of the independent variables on the discriminant function than the standardized canonical coefficients (Hair et al., 1998).

Table 21

Canonical Discriminant Function Coefficients (Standardized Coefficients) Function 1

Variable	2005 Coefficients*	2005 Rank	2004 Coefficients*	2004 Rank
PEER	0.902	1	1.066	1
RECRUT	0.133	7	-0.058	9
QUANT	0.077	9	0.270	3
ANALYT	0.168	6	-0.233	4
ACCEPT	-0.072	10	-0.175	5
PHDFAC	0.229	4	0.116	8
MSFAC	0.096	8	not available	11
NAE	-0.025	11	-0.041	10
RSCHEX	0.183	5	0.164	6
RSCFAC	0.684	2	0.544	2
PHDNO	-0.279	3	-0.134	7

*2005 and 2004 *U.S. News* ranking data used for engineering colleges offering graduate programs.

It is interesting to note that, of the top three variables that most impacted the discriminant function as revealed by the structure matrix coefficients and canonical discriminant function coefficients, peer assessment score was the only common variable with the most impact. The other two variables in the structure matrix table (Table 20) were the recruiter assessment score variable and the research expenditure variable, whereas the other variable in the canonical discriminant function table (Table 21) was research expenditures per faculty. Thus, one can conclude that reputation and level of research funding are the main indicators of the *U.S. News* ranking algorithm.

The second research question asked what phenomena influenced peer assessment score; that is, as peers were completing their *U.S. News* surveys, what phenomena were likely to influence their ratings of institutions. A common sense observation of the *U.S. News* ranking methodology (Table 1 in Chapter 3) revealed that peer score was likely to be the most influential variable in the prediction models. The sum of the peer and recruiter assessment scores is heavily weighted at .40 in the *U.S. News* ranking algorithm, and within this quality assessment rating, peer assessment score is weighted at .25 and recruiter assessment score at .15. A further test of the predominant influence of peer assessment score in the *U.S. News* ranking algorithm is demonstrated in Table 22. Using peer score as the only indicator to classify institutions within their original tier membership with the 2005 *U.S. News* data set in the discriminant analysis test, peer score was able to correctly classify cases with a 77% accuracy upon cross-validation (Table 22) compared with a correct classification rate of 81.9% (Table 12) using all 11 *U.S. News* indicators.

Table 22

Cross-Validated Tier Classification Results Peer Assessment Score 2005 *U.S. News* Data

Yr: 2005		Predicted Group Membership				
Type	*Tier*	*1*	*2*	*3*	*4*	*Totals*
Count	*1*	*15*	6	0	0	21
	2	0	*17*	4	1	22
	3	0	5	*15*	2	22
	4	0	0	2	*20*	22
Percentage	*1*	*71.4%*	28.6%	0.0%	0.0%	100%
	2	0.0%	*77.3%*	18.2%	4.5%	100%
	3	0.0%	22.7%	*68.2%*	9.1%	100%
	4	0.0%	0.0%	9.1%	*90.9%*	100%

77.0% of cross-validated grouped cases correctly classified. N=83; 4 cases with missing data were excluded.

As demonstrated by the *U.S. News* ranking algorithm (Table 1 in Chapter 3), the structure coefficients of the *U.S. News* indicators (Table 20), and the cross-validated classification hit ratio (Table 22), reputational assessment by peers has the strongest influence in the ranking models. Further analyses were conducted to determine which factors influenced peer assessment. Research studies have suggested that a “halo effect” exists between peer assessment ratings of program quality and program size (National Research Council, 2004; Astin, 1992) and the reputation of a few well-known faculty (Stuart, 1995; Rogers & Rogers, 1997; Fombrun, 1996) also known as a “star effect”

(National Research Council, 2004). If these phenomena held true, then there would be a high correlation between peer assessment score and the number of full-time graduate students (halo effect) and the number of full-time engineering faculty who were members of the prestigious National Academy of Engineering (star effect), all variables in the *U.S. News* engineering methodology.

To test the "halo effect" assumption, Pearson correlation between the 2005 *U.S. News* peer assessment scores and the number of full-time master's and PhD students resulted in a significant correlation of .695. Similarly, to test the "star effect" assumption, Pearson correlation between 2005 *U.S. News* peer assessment scores and star faculty, the number of faculty who are members of the prestigious National Academy of Engineering, revealed a significant correlation of .677 (Table 6). High program enrollment and the reputation of nationally renowned faculty are likely to influence the perception of program quality in reputational ratings by peers in the *U.S. News* ranking methodology. The hypothesis holds true: the profile of a successful engineering college could then be described as one that already had a strong academic reputation (usually established, elite universities), large graduate student enrollment, many faculty with membership in the National Academy of Engineering, and high levels of research funding. Thus for institutions with dreams of being ranked with the top tier institutions, they would have to improve in these areas.

The third research question asked how accurately the pBAP model and the replicate *U.S. News* model were able to predict the following year's *U.S. News* rankings. The pBAP model and the *U.S. News* replicate model were able to predict tier movement with approximately equal accuracy. The pBAP model predicted 88.9% of the institutions would remain in the same ranking tier in the following year and the *U.S. News* replicate model 86.4%, compared with 87.7% of the institutions that actually remained in the same ranking tier (Table 19). Tier movement from one year to the next was quite small: 11.1% for the pBAP model and 13.6% for the replicate *U.S. News* model, compared with 12.3% actual tier movement. However, the pBAP model is easier to calculate and use for tier prediction since the overall discriminant score for each case only involves substituting variable values into the linear equation, whereas the *U.S. News* model also involves recalculating the mean and standard deviation of each independent variable.

CHAPTER 5

SUMMARY OF FINDINGS AND RECOMMENDATIONS FOR FUTURE RESEARCH

Introduction

College rankings have become a lucrative and competitive business (Pike, 2004; Walpole, 2003; Hunter, 1995; Dichev, 2001; McDonough et al., 1998; Webster, 1992a). It is common practice for universities not ranked in the top tier in national rankings to develop strategies for improving their standing (Arnone, 2003; Kleiner, 2004; Stecklow, 1995; Karl, 1999; Mallette, 1995; Dichev, 2001; Hunter, 1995; Thompson, 2000; Hossler, 2000; Ridley et al., 2001). Two predictive models were created in this study as a tool for institutional researchers: (a) a *U.S. News* replicate model to predict discrete changes in position order in the rankings, and (b) a tier model (pBAP) using the *U.S. News* ranking data to predict tier movement in the rankings. In this chapter, a summary of the study design, findings from the study, recommendations, and implications for future research are presented.

Statement of the Problem and Purpose of the Study

Improving national ranking is an increasingly important issue for university administrators (Arnone, 2003; Kleiner, 2004; Stecklow, 1995; Karl, 1999; Mallette, 1995;

Dichev, 2001; Hunter, 1995; Thompson, 2000; Hossler, 2000; Ridley et al., 2001). While research has been conducted on performance measures in higher education, research designs have lacked a predictive quality. According to Doerfel and Ruben (2002):

> As with business, higher education indicators have tended to be primarily historical, limited in predictive power, often incapable of alerting institutions to changes in time to respond, and lacking adequate attention to important but difficult-to-quantify dimensions. Ironically, the emphasis on easy-to-quantify, limited measures has, in a manner of speaking, "come to haunt" in the form of popularized college rating systems with which educators generally are frustrated and critical. But these are used consistently as the measures against which universities are evaluated by their constituents. (p. 20)

The purpose of this study was to develop a predictive model for benchmarking academic programs (pBAP) for engineering colleges offering graduate programs based on the *U.S. News* college ranking methodology. This study was not concerned with the validity of the *U.S. News* methodology, but rather with the development of a predictive model within the existing *U.S. News* methodology. The pBAP model can be used to predict what changes in values would cause a college to move up or down a tier in ranking over time. A replicate study of the *U.S. News* methodology was also developed in this study to predict discrete changes in a college's ranking from year to year.

Methodology and Data Collection

Classification discriminant analysis (DA) was deemed the appropriate statistical test to develop the pBAP model. DA identifies which variable means differentiate between predefined groups and predicts group membership of changed variable values or new cases (StatSoft Inc., 2004). A variate, a linear equation of weighted variables, is summed to calculate a discriminant score for each case to distinguish group membership

(Hair et al., 1998). SPSS software was used to analyze the data. This study also included replication of the *U.S. News* ranking methodology for engineering colleges offering graduate programs to predict discrete changes in ranking. Together, the pBAP tier model and the *U.S. News* replication model provide a tool to predict tier ranking (longitudinal goal) and discrete ranking (short-term goal). The predictive model developed in this study was based on the 2005 *U.S. News* data for engineering colleges offering graduate programs. To test the predictive accuracy of the pBAP model, the 2005 *U.S. News* data were entered into the pBAP variate developed using the 2004 *U.S. News* data. The 2003 *U.S. News* data only contained rankings for 50 engineering colleges, an insufficient amount of cases to be used in the pBAP model. The 2003 *U.S. News* data set was used to examine the stability of the 20 top-ranked engineering colleges over a three-year period (2003–2005) and not in the predictive model. The data sets were collected by purchasing online premium access to the *U.S. News* Web site and stored each year in an Excel file by the researcher. The *U.S. News* variables and ranking methodology are listed in Table 1 in Chapter 3.

Data Analysis

The success of the *U.S. News* methodology replication study and its ability to predict the following year's rankings were presented in Chapter 4. The steps used to create the pBAP equation (see Equation 6 in Chapter 4) and the ability of the pBAP model to predict the following year's tier placement in the rankings (see Table 19 in Chapter 4) were demonstrated. The three research questions were addressed.

Summary of Findings

The ability to use the current model to predict tier movement in the future, using the 2004 pBAP model and the 2004 *U.S. News* replicate model to predict institutions that demonstrated tier movement in the 2005 *U.S. News* rankings, was presented in Table 19 in Chapter 4. Only few institutions actually demonstrated tier movement in the following year; most of the institutions remained within the same tier. Tier movement from one year to the next was quite small: 11.1% for the pBAP model and 13.6% for the replicate *U.S. News* model, compared with 12.3% actual tier movement. This finding was consistent with a study by Ridley et al. (2001) in which they examined the transition between tiers using the *U.S. News* ranking data for six years (1996–2001) for 162 four-year liberal arts colleges and found that most of the institutions remained in the original tier in a year-to-year comparison; only 14% of the institutions showed tier movement over the six-year period. Nevertheless, Ridley et al. (2001) recommended that institutions trying to improve their standing in the rankings should not be discouraged from making improvements that would be beneficial to the institution. Moreover, it was important for institutions to not only know their tier ranking but also their position within the tier to determine the distance to the next border, and especially around tier borders that tended to show the most volatility (Ridley et al., 2001).

The pBAP model and the *U.S. News* replicate model were able to predict tier movement with about equal accuracy. However, the pBAP model is easier to calculate and use for tier prediction since the overall discriminant score for each case only involves substituting variable values into the linear equation, whereas the *U.S. News* model also involves recalculating the mean and standard deviation of each independent variable.

Conclusions of the Study and Recommendations

The first research question asked which *U.S. News* variables contributed most to the prediction model variate to discriminate between rankings or tiers. Tsakalis and Palais (2004) analyzed the 2002 *U.S. News* rankings of engineering colleges offering graduate programs and concluded that three variables could be efficiently manipulated for the most impact on the rankings: (a) the number of faculty members in the National Academy of Engineering (NAE), (b) the number of PhD degrees awarded, and (c) the amount of research expenditures. For the engineering college used as an example in their study, they recommended a strategy of "hiring one research-active faculty [at] $150k/year, hiring one NAE faculty [at] $300k/year, hiring one PhD student [at] $45k/year, and [a] 10:1 return-on-investment for attracting new sponsored projects" (Tskalis & Palais, 2004, p. 262). In this study, however, of the 11 variables in the *U.S. News* model, the likelihood of improving an institution's standing in the rankings was greater when increasing the following three variables as revealed by the structure matrix coefficients in Table 20 in Chapter 4: peer assessment score, recruiter assessment score, and research expenditures.

The second research question asked what phenomena influenced peer assessment score; that is, as peers were completing their *U.S. News* surveys, what phenomena were likely to influence their ratings of institutions. Peer assessment score had an almost monopolistic impact on the rankings. Using peer score as the only indicator to classify institutions within their original tier membership with the 2005 *U.S. News* data, peer score was able to correctly classify cases 77% upon cross-validation (Table 22 in Chapter 4) compared with a correct classification rate of 81.9% using all 11 indicators (Table 12 in Chapter 4). Moreover, an engineering college's undergraduate ranking was likely to have a strong correlation with its graduate rankings because the undergraduate rankings were calculated solely on one indicator, peer assessment score. This study found that an engineering college's 2003 and 2004 *U.S. News* undergraduate rankings had a strong correlation (greater than .90) with their 2004 and 2005 *U.S. News* graduate rankings, respectively.

Research studies have suggested that a "halo effect" exists between peer assessment ratings of program quality and program size (National Research Council, 2004; Astin, 1992) and the reputation of a few well-known faculty (Stuart, 1995; Rogers & Rogers, 1997; Fombrun, 1996) also known as a "star effect" (National Research Council, 2004). Vojak et al. (2003) analyzed the relationship between *U.S. News* reputation rankings for engineering colleges and departments for 13 years and found that engineering departments with the largest number of students and faculty were more likely to generate academic leaders and to bring visibility (renown) to the college (the "halo effect" phenomenon). To determine what phenomena impacted peer assessment scores in this study, correlation tests were run. Pearson correlation between the 2005 *U.S. News* peer assessment scores and the number of full-time master's and PhD students resulted in a significant correlation of .695. Similarly, to test the "star effect" assumption, Pearson correlation between 2005 *U.S. News* peer assessment scores and star faculty, the number of faculty who are members of the prestigious National Academy of Engineering, revealed a significant correlation of .677 (Table 6 in Chapter 4). High program enrollment and the reputation of nationally and internationally renowned faculty are likely to influence the perception of program quality in reputational ratings by peers in the *U.S. News* ranking methodology.

The profile of a successful engineering college could then be described as one that had a strong academic reputation (usually established, elite universities), high graduate student enrollment, many faculty with membership in the National Academy of Engineering, and faculty productive in bringing in research funding. It is recommended that an emphasis on improving reputation by influencing peer assessment scores is a more efficient and cost-effective strategy to increase ranking status than the recommendations of Tsakalis and Palais described at the beginning of this section. If one followed the recommendations of Tsakalis and Palais (2004), it would cost the engineering college a minimum of $495,000 annually in salary for those three hires. Whereas, for example, two half-page advertisements in the *U.S. News* guidebooks and two full-page advertisements in the college ranking editions of *U.S. News* magazine would cost in the neighborhood of $100,000 annually, depending on the target market selected for the magazine ads and not including discounts for multiple ads (Jim Robben, EMI Network, personal communication, April 15, 2004).

According to Fombrun (1996, p. 6), “The proliferation of such subjective rankings as ‘best managed,’ ‘most innovative,’ and ‘most admired’ attests to the growing popularity of reputation as a tool for assessing companies.” Academic quality rankings can affect the size and quality of an institution’s applicant pool (Webster, 1992a). Monks and Ehrenberg (1999) examined the impact of *U.S. News* rankings on students’ college choices and found that institutions that experienced a decrease in the rankings were likely to attract fewer students, thereby forcing the institution to admit more applicants (increasing their selectivity rate) and fewer of those students tended to enroll (causing a decrease in yield). McDonough et al. (1998) found that of the students who consulted college rankings and believed them to be very important in college choice, 70% believed that academic reputation influenced admittance into a top graduate school. Graduates from colleges and universities with renowned academic reputation are more likely to have an edge in the job market for high-paying jobs or admission into top graduate schools over graduates from lesser-known institutions (Morse & Gilbert, 1995; Ehrenberg, 2003). In fact, Americans are fascinated by rankings: one can find rankings of sports teams, tall buildings, hospitals, and cities with the best business climate, among other things (Hossler, 2000). And the media continues to perpetuate the image of the elite and highly visible Ivy League colleges, but for the "equally pricey but not-so-well-knowns" (Moll & Wright, 1998, p. 159), one rarely reads about them in the media. According to Moll and Wright (1998):

> …these are the colleges that, lacking visibility and therefore positive image, are scrambling for students, worrying about net tuition revenue, and offering fire-sale prices to fill the beds. . . (p. 159) . . .One theory is that the media know that Americans are tremendously status conscious, with advertising and brand name a major aspect of capitalism. (p. 157)

Like private enterprises, investing in an effective marketing campaign is likely to increase favorable perception of quality program offerings and disseminate information on the engineering college to a broader audience. A marketing campaign would be especially effective for an engineering college that is being held back in national prominence because of the low ranking of the university overall. Colleges of Engineering and Computer Science are among the target disciplines receiving state and national attention. A recent article by Adam Segal (2004) suggests that America is fast losing its innovative edge in the development of new technologies and industries to global competition, in particular, from Asia. An effective advertising campaign to influence peer and public perception can be compared to mounting an effective election campaign. The candidate publicizes his or her platform through the advertising campaign to influence votes. Business schools like the University of Chicago, the University of Maryland, and New York University hired public relations firms to develop marketing strategies when *Business Week* began ranking business schools across the nation based on reputational scores of former students and industry recruiters (Fombrun, 1996). Savvy colleges and universities employ sophisticated marketing and business tools, ranging from videos, telemarketing, targeted mailings, college ranking publications, college guidebooks, to television, radio, and billboard advertisements, to build a reputation and attract students (Hossler & Foley, 1995; McGuire, 1995). Colleges and universities can "make themselves hot with some savvy self-promotion" (Shea, 2004b, p. 57), and a new wave of self-promotion can be seen on reality TV shows about student life on campus (Shea, 2004b). In short, a good reputation increases the credibility of the organization in the eyes of the public (Fombrun, 1996).

The third research question asked how accurately the pBAP model and the replicate *U.S. News* model were able to predict the following year's *U.S. News* rankings. As discussed in an earlier section, *Summary of Findings,* tier movement from one year to the next was quite small; nevertheless, Ridley et al. (2001) recommended that institutions trying to improve their standing in the rankings should not be discouraged from making improvements that would be beneficial to the institution. The pBAP model aims to provide a tool to assist institutional researchers and university administrators to predict future ranking status in the *U.S. News* college rankings by identifying areas for institutional development by comparing internal practices with best practices elsewhere. According to Robert Morse, Director of Data Research, *U.S. News,* changes are made to its ranking algorithms periodically in response to suggestions for improvement from the higher education community (Morse & Gilbert, 1995). Clarke (2001) compared the 1995 and 2000 *U.S. News* graduate rankings for engineering and found a .88 correlation, taking into account changes to the methodology over a period of six years. As a precaution, however, the discriminant analysis procedure should be rerun each year for new weights for the pBAP model for the following year's predictions. A Web-interactive rank calculator of the pBAP model (completed) was created as part of this study as a user-friendly tool for the deans of engineering colleges to determine the most effective strategy of achieving long-term goals to improve in ranking status. The development of a Web-interactive rank calculator of the replicate *U.S. News* model is in progress to determine the most effective strategy of achieving short-term goals to improve standing in the rankings.

As the researcher was finalizing this dissertation, the 2006 *U.S. News & World Report, America's Best Graduate Schools* was published. *U.S. News* had made changes to the indicators and the indicator weights in the algorithm for the 2006 rankings. Table 23 compares the 2005 and 2006 *U.S. News* graduate engineering ranking methodology. However, comparison of the structure matrix coefficients for of 2005 and 2006 *U.S. News* data revealed that the contribution of the independent variables to the discriminant function remained the same, with peer assessment score, recruiter assessment score, and research expenditures as the top three variables with the most impact (Table 24). The cross-validated classification accuracy of cases into the groups in the pBAP model using the 2006 *U.S. News* data declined to 73.9% compared with 81.9% for the 2005 *U.S. News* data.

Table 23

Comparison of Weights 2005 and 2006 *U.S. News* Graduate Engineering Methodology

2005 & 2006 Dimensions	2005 Indicators	2006 Indicators
Quality Assessment (.40)	Peer (.25)	Peer (.25)
	Recruiter (.15)	Recruiter (.15)
Student Selectivity (.10)	Quantitative GRE (.45)	Quantitative GRE (.0675)
	Analytical GRE (.45)	Analytical GRE (removed)
	Acceptance (.10)	Acceptance (.0325)
Faculty Resources (.25)	PhD students/faculty (.30)	PhD students/faculty (.075)
	MS students/faculty (.15)	MS students/faculty (.0375)
	Faculty in NAE (.30)	Faculty in NAE (.075)
	PhD degrees (.25)	PhD degrees (.0625)
Research Activity (.25)	Research expenditures (.60)	Research Expenditures (.15)
	Research per faculty (.40)	Research per faculty (.10)
Total indicators	11	10

Overall rank: Data were standardized about their means, and standardized scores were weighted, totaled, and rescaled so that the top-scoring school received 100; others received their percentage of the top score.

See Table 1 in Chapter 3 for full description of indicators. Source: www.usnews.com

Table 24

Comparison of Structure Matrix Coefficients for the 2005 and 2006 *U.S. News* Data

2005 Indicator	2005 Function 1	2005 Contribution Rank Order	2006 Indicator	2006 Function 1	2006 Contribution Rank Order
PEER05	.660	1	PEER06	.715	1
RECRUT05	.592	2	RECRUT06	.600	2
PHDFAC05	.372	4	PHDFAC06	.400	4
NAE05	.323	6	NAE06	.366	6
RSCFAC05	.303	7	RSCFAC06	.350	7
ACCEPT05	-.140	10	ACCEPT06	-.176	9
RSCHEX05	.382	3	RSCHEX06	.432	3
PHDNO05	.350	5	PHDNO06	.391	5
MSFAC05	-.014	11	MSFAC06	-.002	10
ANALYT05	.249	8	ANALYT06	removed	
QUANT05	.240	9	QUANT06	.242	8

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions.

Nevertheless, the 73.9% classification rate still showed significant improvement over a 25% increase in the proportional chance accuracy rate achievable by chance alone of 31.3%. The Press' *Q* statistic (for N = 88, n = 65, and K = 4) was 112, which was significantly greater than the critical value $\chi^2_{(df=3,\ \alpha=.01)} = 11.35$. Therefore, the discriminatory power of the pBAP model to correctly classify cases into the groups is significantly better than that of chance, thereby supporting cross-validation. The critical value of χ^2 was retrieved from the table of critical values on the Web site of the National Institute of Standards (retrieved April 9, 2005). The Box's *M* statistic remained statistically significant at p<.000, still denoting violation of homogeneity even with the change in variable weights and the removal of one variable (average analytical GRE score) in the 2006 *U.S. News* ranking model.

The peer and recruiter score of the test case institution used to demonstrate the pBAP model in Chapter 4 each increased by 0.2 each (from 2.2 to 2.4 and 2.6 to 2.8 respectively). As observed in Table 6, the correlation between peer and recruiter score was .963. Thus, if peer score increased or decreased, it was likely that recruiter score would manifest the same magnitude and direction as the peer score and vice versa. A correlational analysis comparing the ranking of the top 25 engineering colleges with the engineering specialty rankings revealed that all but one engineering college had a ranked electrical, computer, and mechanical engineering department. In addition, the electrical, computer, and mechanical engineering specialties correlate the highest with college rankings. This finding is not unusual, given that these programs tend to attract students and therefore have the highest enrollment. Program size contributes to a "halo effect," where large program size correlates highly with peer assessment score.

As mentioned above, a study by Clarke (2001) compared the 1995 and 2000 *U.S. News* graduate rankings for engineering and found a .88 correlation, taking into account changes to the methodology, thereby demonstrating that changes to the engineering methodology did not have a significant impact on the rankings. Based on the brief analyses above of the 2006 *U.S. News* rankings, changes in the methodology are not likely to greatly impact the predictive power of the pBAP model.

Implications for Future Research

As the most well-known source for college rankings (Stuart, 1995), the *U.S. News* college rankings play an important role in students' choice of colleges, legislators' funding allocation decisions, and the public's perception of tax dollar expenditures on institutions of higher education. Many public universities have a goal to move up in the national rankings as part of their strategic plans (Arnone, 2003). Although the data used in this study were specifically for graduate engineering programs at the college level, the methodology can be adapted to predict future *U.S. News* rankings for other graduate disciplines (business, education, law, and medicine) and at the university level.

APPENDIX A

CRONBACH'S ALPHA 2005 *U.S. NEWS* DATA

Reliability

Method 2 (covariance matrix) will be used for this analysis ******

R E L I A B I L I T Y   A N A L Y S I S   -   S C A L E   (A L P H A)

		Mean	Std Dev	Cases
1.	PEER05	3.2687	.6761	83.0
2.	RECRUT05	3.3651	.5047	83.0
3.	QUANT05	752.3494	19.7314	83.0
4.	ANALYT05	692.8313	33.7736	83.0
5.	ACCEPT05	.2546	.1289	83.0
6.	PHDFAC05	2.9639	1.1719	83.0
7.	NAE05	.0424	.0432	83.0
8.	RSCHEX05	57.5096	49.3300	83.0
9.	RSCFAC05	415.3373	190.3591	83.0
10.	PHDNO05	60.3373	49.1693	83.0
11.	MSFAC05	2.6390	1.9208	83.0

Correlation Matrix

	PEER05	RECRUT05	QUANT05	ANALYT05	ACCEPT05
PEER05	1.0000				
RECRUT05	.9639	1.0000			
QUANT05	.5644	.5325	1.0000		
ANALYT05	.6066	.6167	.6866	1.0000	
ACCEPT05	-.3285	-.3086	-.5109	-.4705	1.0000
PHDFAC05	.7082	.6696	.5110	.4756	-.4699
NAE05	.6805	.6980	.4971	.5385	-.2393
RSCHEX05	.7628	.7116	.3614	.3069	-.1758
RSCFAC05	.4692	.4757	.3220	.2681	-.1830
PHDNO05	.8357	.7763	.3553	.3181	-.1089
MSFAC05	-.0927	-.1441	-.0522	-.3336	.5109

	PHDFAC05	NAE05	RSCHEX05	RSCFAC05	PHDNO05
PHDFAC05	1.0000				
NAE05	.6904	1.0000			
RSCHEX05	.5247	.4199	1.0000		
RSCFAC05	.5160	.5316	.6111	1.0000	
PHDNO05	.5526	.4876	.8883	.4243	1.0000
MSFAC05	-.0686	-.0276	.0974	.1671	.1238

	MSFAC05
MSFAC05	1.0000

R E L I A B I L I T Y   A N A L Y S I S   -   S C A L E   (A L P H A)

N of Cases = 83.0

Reliability Coefficients 11 items

Alpha = .4958 Standardized item alpha = .8397

APPENDIX B

CRONBACH'S ALPHA 2004 *U.S. NEWS* DATA

Reliability

Method 2 (covariance matrix) will be used for this analysis

R E L I A B I L I T Y   A N A L Y S I S   -   S C A L E   (A L P H A)

		Mean	Std Dev	Cases
1.	PEER04	3.3593	.6473	81.0
2.	RECRUT04	3.5185	.4642	81.0
3.	QUANT04	752.0000	19.5666	81.0
4.	ANALYT04	689.4815	27.9652	81.0
5.	ACCEPT04	.2999	.1485	81.0
6.	PHDFAC04	2.7914	1.1348	81.0
7.	NAE04	.0413	.0446	81.0
8.	RSCHEX04	54.3494	44.9659	81.0
9.	RSCFAC04	404.2778	186.7326	81.0
10.	PHDNO04	61.1481	49.5704	81.0

Correlation Matrix

	PEER04	RECRUT04	QUANT04	ANALYT04	ACCEPT04
PEER04	1.0000				
RECRUT04	.9368	1.0000			
QUANT04	.4686	.4033	1.0000		
ANALYT04	.5040	.4787	.7559	1.0000	
ACCEPT04	-.4886	-.4435	-.5440	-.5943	1.0000
PHDFAC04	.6865	.6015	.4937	.4557	-.5978
NAE04	.7040	.6714	.4352	.4551	-.3879
RSCHEX04	.7388	.6503	.3965	.2656	-.2746
RSCFAC04	.3767	.3166	.3802	.2578	-.2627
PHDNO04	.7954	.7248	.3505	.2596	-.2253

	PHDFAC04	NAE04	RSCHEX04	RSCFAC04	PHDNO04
PHDFAC04	1.0000				
NAE04	.6736	1.0000			
RSCHEX04	.4771	.4161	1.0000		
RSCFAC04	.4635	.4895	.5823	1.0000	
PHDNO04	.4802	.4143	.8710	.2821	1.0000

N of Cases = 81.0

R E L I A B I L I T Y   A N A L Y S I S   -   S C A L E   (A L P H A)

Reliability Coefficients   10 items

Alpha = .4601   Standardized item alpha = .8319

LIST OF REFERENCES

Altbach, P. G. (2001, Spring). Higher education and the WTO: Globalization run amok. *International Higher Education.* Reprinted in revised form from the *Chronicle of Higher Education* by the Center for International Higher Education, Boston College. Retrieved January 19, 2003, from http://www.bc.edu/bc_org/Avp/soe/cihe/newsletter/News23/text0001.htm.

Altbach, P. G. (1998). Patterns in higher education development. *Comparative higher education: Knowledge, the university, and development* (pp. 15, 29). Greenwich, CT: Ablex Publishing.

Arnone, M. (2003, January 3). The wannabes. *The Chronicle of Higher Education,* 18–20.

Arnone, M., Hebel, S., Schmidt, P. (2003, January 3). Another bleak budget year. *The Chronicle of Higher Education,* A21–A22.

Astin, A. W. (1992, November 1). Responses to the rankings. *Change 24(6).* Retrieved November 5, 2004, from http://bll.epnet.com, Item #0009-1383.

Barak, R. J., Kniker, C. R. (2002, Summer). Benchmarking by state higher education boards. *New Directions for Higher Education 118*, 93–101.

Bracey, G. W. (2003). *What you should know about the war against America's public schools.* Boston, MA: Allyn and Bacon.

Brainard, J., & Burd, S. (2004, June 11). College officials brace for federal spending cuts in fiscal 2006 [Electronic version]. *Chronicle of Higher Education 50(40),* A16. Retrieved July 19, 2004, from http://chronicle.com/weekly/v50/i40/40a1602.htm.

Breton, G. (2002, September 18). Higher education: From internationalization to globalization. *Globalisation: What issues are at stake for universities?* Proceedings of the conference. *International Association of Universities.* Retrieved January 19, 2003, from http://www.ulaval.ca/BI/Globalisation-Universities/pages/actes/Gilles-Breton-ang.pdf.

Brown, J. D. (2002, February). The Cronbach alpha reliability estimate. *Shiken: JALT Testing & Evaluation SIG Newsletter 6(1).* Retrieved July 11, 2004, from http://www.jalt.org/test/bro_13.htm.

Brunner, B. (n.d.). Bakke and beyond: A history and timeline of affirmative action. *Infoplease.com.* Pearson Education. Retrieved February 6, 2005, from http://www.infoplease.com/spot/affirmative1.html.

Cary, P. (2000, August 23). *U.S. News and World Report* response to Nicholas Thompson's article, Playing with numbers: How U.S. News mismeasures higher education and what we can do about it [Letter to the editor]. *Washington Monthly.* Retrieved August 9, 2004, from http://www.washingtonmonthly.com/features/2000/0009.thompson.html.

Cary, P., & Duffy, B. (1999, September 7). Dissension in the rankings: *U.S. News* responds to *Slate's* "best colleges" story. *Slate.* Retrieved August 9, 2004, from http://slate.msn.com/id/34278.

Clarke, M. (2001, April). Quantifying quality: Technical issues surrounding "U.S. News and World Report's" rankings of U.S. colleges and graduate schools. Paper presented at the Annual Meeting of the American Educational Research Association (Seattle, WA, April 10–14, 2001). Retrieved November 2, 2004, from http://ww.eric.ed.gov, ERIC# ED454257.

College Board (2004, October 19). The College Board's annual reports on trends in college pricing 2004. Retrieved December 17, 2004, from http://www.collegeboard.com/press/article/0,,3899,00.html.

Dichev, I. (2001). News or noise? Estimating the noise in the *U.S. News* university rankings [Electronic version]. *Research in Higher Education, 42(3),* 237–266. Retrieved July 19, 2003, from http://www.kluweronline.com/issn/0361-0365/contents.

Dillon, S. (2004, December 22). America slipping as education hub: Indians and Chinese taking GRE down by half. *New York Times News Service* posted on *The Indian Express.* Retrieved December 22, 2004, from http://www.indianexpress.com.

Doerfel, M. L., & Ruben, B. D. (2002, Summer). Developing more adaptive, innovative, and interactive organizations. *New Directions for Higher Education 118*, 5–24.

Drew, D. E., & Karpf, R. (1981). Ranking academic departments: Empirical findings and a theoretical perspective. *Research in Higher Education 14(4),* 305–320.

Drucker, P. F. (1969). Has success spoiled the schools? *The age of discontinuity* (p. 311). New York: Harper & Row.

Ehrenberg, R. G. (2003). Reaching for the brass ring: How the "U.S. News & World Report" rankings shape the competitive environment in U.S. higher education [Electronic verson]. *The Review of Higher Education 26(2),* 145–162. Retrieved February 4, 2004, from: http://muse.jhu/journals/review_of_higher_education/ V026/26.2ehrenberg.html or http://ejournals.ebsco.com/direct.asp? Article ID=LVX18H2HD3T9PR8PVGQ9.

Elfin, M. (1992, November 1). Responses to the rankings. *Change 24(6).* Retrieved November 5, 2004, from: http://bll.epnet.com, Item# 0009-1383.

El Khawas, E. (1992, November 1). Response to the rankings. *Change 24(6).* Retrieved November 5, 2004, from http://bll.epnet.com, Item# 0009-1383.

Ewers, J. (2004, August 30). Getting into college: What you need to know now about the admission process. *U.S. News & World Report 137(6)*, 68–74.

Fombrun, C. J. (1996). *Reputation: Realizing value from the corporate image.* Boston, MA: Harvard Business School Press.

The Futures Project. (2000, October). The universal impact of competition and globalization in higher education. *The Futures Project: Policy for higher education in a changing world.* Retrieved January 19, 2003, from http://www. futuresproject.org/publications/universal_impact.pdf.

Garson, G. D. (n.d.). Scales and standard measures. In *PA 765 Statnotes: An online textbook.* Retrieved July 9, 2004 from http://www2.chass.ncsu.edu/garson/ pa765/standard.htm.

Gorman, L. (2003, December). State education subsidies shift students to public universities. *National Bureau of Economic Research (NBER) Digest Online.* Retrieved July 17, 2004, from http://www.nber.org/digest/dec03/w9720.html.

Gottlieb, B. (1999, Sept. 1). Cooking the books: How *U.S. News* cheats in picking its "best American colleges." *Slate.* Retrieved August 9, 2004, from http://slate.msn.com/id/34027/.

Griffin, D. (1999, Winter). Education not an island. *State Education Leader 17*(1). Retrieved April 1, 2003, from http://ecs.org/Clearinghouse/33/56/3356.htm.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). Multiple discriminant analysis and logistic regression. *Multivariate data analysis* (5th ed.) (pp. 239–325). Upper Saddle River, NJ: Prentice-Hall.

Henry, M., Lingard, B., Rizvi, F., & Taylor, S. (2001). *The OECD, globalization and education policy* (pp. 52–53). Oxford, England: Elsevier Science.

Hossler, D. (2000, March–April). The problem with college rankings. *About Campus 5(1),* 20–24.

Hossler, D. (1998). Everybody wants to be number one: The effects of the media's college rankings. In Gene I. Maeroff (Ed.), *Imaging education: The media and schools in America* (pp. 161–177). New York, NY: Teachers College Press.

Hossler, D., & Foley, E. M. (1995, Winter). Reducing the noise in the college choice process: The use of college guidebook and ratings. In R. Dan Walleri and Marsha K. Moss (Eds.), *Evaluating and responding to college guidebooks and rankings. New Directions for Institutional Research 88,* 21–30. San Francisco: Jossey-Bass.

Hubbell, L. W. L., Massa, R. J., Lapovsky, L. (2002, Summer). Using benchmarking to influence tuition and fee decisions. *New Directions for Higher Education 118*, 39–63.

Hunter, B. (1995, Winter). College guidebooks: Background and development. In R. Dan Walleri and Marsha K. Moss (Eds.), *Evaluating and responding to college guidebooks and rankings. New Directions for Institutional Research 88,* 5–12. San Francisco: Jossey-Bass.

Ince, M. (2004a, November 5). Wealthy set to challenge US while others languish. *The Times Higher Education Supplement,* 8–9.

Ince, M. (2004b, November 5). West is best but there's a rich feast in the East. *The Times Higher Education Supplement,* 12–13.

International Institute of Education. (2001, November 13). Record total of 547,867 international students on U.S. campuses enrollment rises 6.4%: largest increase since 1980. In *Open Doors 2001*. Retrieved August 25, 2002, from http://www.iie.org/Content/ContentGroups/Announcements/Record_total_of_547,867_international_students_on_U_S_campuses_enrollment_rises_6_4%;_largest_incre.htm.

Karl, J. (1999, Sept. 10). Taste: The great campus celebrity contest—Do those ballyhooed college rankings mean anything at all? A tutorial [Electronic Version]. *The Wall Street Journal* (Eastern edition), New York, NY, W.13. Retrieved September 20, 2004 from http://proquest.umi.com/pqdlink?index=1&sid=1&srchmode=1&vinst=PROD&fmt=3&startpa.

Karlöf, B., & Östblom, S. (1993). *Benchmarking: A signpost to excellence in quality and productivity* (pp. 1–2). West Sussex, England: John Wiley & Sons.

Kleiner, C. (2004). Decades of ranking. *America's best colleges,* 74–76.

Mallette, B. I. (1995, Winter). *Money Magazine, U.S. News & World Report,* and Steve Martin: What do they have in common? In R. Dan Walleri and Marsha K. Moss (Eds.), *Evaluating and responding to college guidebooks and rankings. New Directions for Institutional Research 88,* 31–43. San Francisco: Jossey-Bass.

McDonough, P. M. (1994, Jul.–Aug.). Buying and selling higher education: The social construction of the college applicant [Electronic Version]. *The Journal of Higher Education 65*(4): 427–446. Retrieved September 17, 2004, from http://links.jstor. Org/sici?sici=0022-1546%28199407%2F08%2965%3A4%3C427%3ABASHET %3E2.0.CO%3B2-H.

McDonough, P. M., Lising Antonio, A., Walpole, M., & Xóchitl Pérez, L. (1998). College rankings: Democratized college knowledge for whom? [Electronic version.] *Research in Higher Education 39(5),* 513–537. Retrieved July 19, 2004, from http://www.kluweronline.com/issn/0361-0365/contents.

McGuire, M. D. (1995, Winter). Validity issues for reputational studies. In R. Dan Walleri and Marsha K. Moss (Eds.), *Evaluating and responding to college guidebooks and rankings. New Directions for Institutional Research 88,* 45–59. San Francisco: Jossey-Bass.

Moll, R. W., & Wright, B. A. (1998). What college selectivity looks like to the public. In Gene I. Maeroff (Ed.), *Imaging education: The media and schools in America.* New York, NY: Teachers College Press.

Monks, J., & Ehrenberg, R. G. (1999, November). U.S. News & World Report's college rankings: Why they do matter. *Change*, 43–51. Retrieved February 12, 2005, from http://www.findarticles.com/p/articles/mi m1254/is 6 31/ai 58178197/print.

Morse, R. J., & Gilbert, J. (1995, Winter). Publishers' perspectives: Magazines. In R. Dan Walleri and Marsha K. Moss (Eds.), *Evaluating and responding to college guidebooks and rankings. New Directions for Institutional Research 88,* 91–107. San Francisco: Jossey-Bass.

Muffo, J. A. (2003). Institutional research support of accountability. In W. E. Knight (Ed.), *The primer for institutional research* (p. 2). Tallahassee, FL: Association for Institutional Research.

National Institute of Standards. (n.d.). Critical values of the Chi-Square distribution. Retrieved April 9, 2005, from: http://www.itl.nist.gov/div898/handbook/ eda/section3/eda3674.htm.

National Opinion Research Center (NORC). (1997). A review of the methodology for the U.S. News & World Report's rankings of undergraduate colleges and universities. Published as part of Nicholas Thompson's research for his piece *Playing with numbers: How U.S. News and World Report mismeasures higher education and the rest of us. Washington Monthly.* Retrieved August 9, 2004, from http:www. Washingtonmonthly.com/features/2000/norc.html.

National Research Council (NRC). (2004). Assessing research-doctorate programs: A methodology study [Electronic version]. In J. P. Ostriker & C. V. Kuh (Eds.). Retrieved May 2004, from http://www.nap.cdn/openbook/030909058X/html/ R1.html.

National Society for Multiple Sclerosis. (2002, Sept. 10). *Research: Measures for use in clinical studies of MS.* Retrieved July 9, 2004, from http://www.nationalmssociety. org/MUCS_glossary3.asp#splithalfreliability.

Neave, G. (n.d.) Globalization: Threat, opportunity, or both? *International Association of Universities.* Retrieved January 19, 2003, from http://www.unesco.org/iau/ Globalization/newsletter-globalization.html.

Newman, F., & Couturier, L. (2001, June). The new competitive arena: Market forces invade the academy. *The Futures Project: Policy for higher education in a changing world.* Retrieved January 19, 2003, from http://www.futuresproject. org/publications/new_competitive_area.pdf.

Novak, R. J. (2002, Summer). Benchmarking distance education. *New Directions for Higher Education 118,* 79–91.

O'Leary, J. (2004, November 5). Top performers on the global stage take a bow. *The Times Higher Education Supplement,* 2–5.

Pike, G. R. (2004, March). Measuring quality: A comparison of *U.S. News* rankings and NSSE benchmarks. *Research in Higher Education 45(2),* 193–206.

Reider, R. (2000). *Benchmarking strategies: A tool for profit improvement* (p. 1). New York: John Wiley & Sons.

Ridley, D. R., Cuevas, N. M., Matveev, A. G. (2001). Transitions between tiers in U.S. News and World Report rankings of colleges and universities. Report retrieved November 2, 2004, from http://www.eric.ed.gov, ERIC# ED457817.

Rogers, E., & Rogers S. J. (1997, May). High science vs. just selling magazines: How the NRC and *U.S. News* graduate rankings compare. *American Association of Higher Education Bulletin 49(9),* 7–10.

Roose, K. D., & Andersen, C. J. (1970). *A rating of graduate programs.* Washington D.C.: American Council on Education.

Rossman, P. (2002). Exploring and creating the future in virtual space. In *The future of higher education: For all worldwide, a holistic view* (vol. 1). Retrieved January 24, 2003, from http://ecolecon.missouri.edu/globalresearch/chapters/1-PREF.html.

Rudolph, F. (1990). The American college & university: A history (pp. 1–22, 201–220). Athens, GA: University of Georgia Press.

Salmi, J. (1999, December 3). Globalisation and the knowledge-based economy pose challenges to all purveyors of tertiary education. *World Bank.* Retrieved January 19, 2003, from http://www.1.worldbank.org/education/tertiary/news_archive/1203d9.html.

Schwab, A. J. (2003, April 2). Strategy for complete discriminant analysis. Class lecture notes powerpoint for sw388r7. Retrieved March 22, 2005, http://www.utexas.edu/Courses/schwab/sw38847/SolvingProblems.

Secor, R. (2002, Summer). Penn State joins the Big Ten and learns to benchmark. *New Directions for Higher Education 118*, 65–77.

Segal, A. (2004, November 17). Is American losing its edge? *YaleGlobal Online*. Copyright 2002–2004 by the Council on Foreign Relations. Reprinted from the November–December 2004 issue of *Foreign Affairs.* Retrieved March 30, 2005, from http://yaleglobal.yale.edu/display.article?id=4893.

Shea, R. H. (2004a). How we got here. *U.S. News & World Report: America's Best Colleges, 2004 Edition,* 70–73.

Shea, R. H. (2004b). How schools get hot. *U.S. News & World Report: America's Best Colleges, 2004 Edition,* 57.

Skarloff, S. (2004). My college, my self. *U.S. News & World Report: America's Best Colleges, 2004 Edition,* 22–23.

SPSS, Inc. (1999). Discriminant analysis. In *SPSS base 10.0 applications guide.* Chicago, IL: Statistical Package for the Social Sciences (SPSS), Inc.

Stanfield, R. L. (1998). The media and public perceptions of tuition costs. In Gene I. Maeroff (Ed.), *Imaging education: The media and schools in America* (pp. 135–146). New York, NY: Teachers College Press.

StatSoft, Inc. (2004). Electronic statistics textbook. Tulsa, OK: StatSoft. Retrieved July 9, 2004, from http://www.statsoft.com/textbook/stdiscan.html.

Stecklow, S. (1995, April 5). Cheat sheets: Colleges inflate SATs and graduation rates in popular guidebooks—Schools say they must fib to U.S. News and others to compete effectively—Moody's requires the truth [Electronic Version]. *The Wall Street Journal* (Eastern edition), New York, NY, A.1. Retrieved September 20, 2004, from http://proquest.umi.com/pdqlink?index=2&sid=5&srchmode=3&vinst =PROD&fmt=3&startpa.

Stuart, D. L. (1995, Winter). Reputational rankings: Background and development. In R. Dan Walleri and Marsha K. Moss (Eds.), *Evaluating and responding to college guidebooks and rankings. New Directions for Institutional Research 88,* 13–20. San Francisco: Jossey-Bass.

Thompson, N. (2000, September). Playing with numbers. *Washington Monthly.* Retrieved August 9, 2004, from http://washingtonmonthly.com/features/ 2000/0009.thompson.html.

Tsakalis, K. S., & Palais, J. C. (2004, July). Improving a school's *U.S. News and World Report* ranking. *Journal of Engineering Education,* 259–263.

U.S. News & World Report (U.S. News). (2003, 2005). *America's best colleges. America's best graduate schools.* Retrieved September 4, 2004 and February 20, 2005, from http://www.usnews.com.

Vojak, B. A., Price, R. L., & Carnahan, J. V. (2003, January). The relationship between department rank and college rank in engineering graduate program rankings conducted by *U.S. News and World Report. Journal of Engineering Education,* 65–71.

Walpole, M. (2003, November 20). How students and families use rankings for college selection. Wisconsin Center for the Advancement of Postsecondary Education (WISCAPE) Forum: *The use and abuse of college rankings.* Retrieved September 20, 2004, from http://www.wiscape.wisc.edu/publications/attachments/434%20 Walpole20Paper.pdf.

Webster, D. S. (1992a, March–April). Are they any good? Rankings of undergraduate education in U.S. News & World Report and Money. *Change 24(2).* Retrieved November 4, 2004, from http://weblinks1.epnet.com, Item# 9207203726.

Webster, D.S. (1992b, September–October). Academic rankings: First on a list of one. *Academe* 19–22.

Webster, D. S. (1986). *Academic quality rankings of American colleges and universities.* Springfield, IL: Charles C. Thomas.

The World Bank. (1994). *Higher education: The lessons of experience.* Washington D.C.: The International Bank for Reconstruction and Development/The World Bank. Retrieved February 1, 2002, from http://wbln0018.worldbank.org/LAC/Lacinfoclient.nsf/d29684951174975c85256735007fef12/8340529a1767bee85256968006861e4?/OpenDocument.

Wright, B. A. (1990–1991, Winter). The rating game: How the media affect college admission. *The College Board Review 158,* 12–17, 31–32.