2009

Punching a Ticket to the Big Dance: A Critical Analysis of At-Large Selection into the NCAA Division I Men's Basketball Tournament

Stephen L. Shapiro
Old Dominion University

Joris Drayer

Brendan Dwyer

Alan L. Morse

Follow this and additional works at: https://digitalcommons.odu.edu/hms_fac_pubs

Part of the Sports Studies Commons

Punching a Ticket to the Big Dance: A Critical Analysis of At-Large Selection Into the NCAA Division I Men's Basketball Tournament

_____

Stephen L. Shapiro
Old Dominion University

Joris Drayer
University of Memphis

Brendan Dwyer
Virginia Commonwealth University

Alan L. Morse
Mississippi State University

_____

The NCAA Division I Men's Basketball Tournament at-large selection process is a highly scrutinized system. A selection committee made up of ten members is asked to select teams based on guidelines set forth by the NCAA. However, committee members indicate that other factors are also considered. Due to the financial benefits associated with a tournament appearance, it is important to closely examine the committee's decision making process in order to identify possible sources of bias, which may put some teams at a competitive disadvantage. The purpose of this study was to examine direct and indirect attributes of teams that finished in the RPI top 100 between 1999 and 2007 (N = 695) to provide a critical analysis of the selection process. The results of a logistic regression model suggested that most of the criticism received by the selection committee is unwarranted. Overall, previous selections matched most of the performance-based guidelines that are used by the selection committee. However, conference classification was a significant predictor of selection, and this may provide evidence of a major conference bias. Specific misclassifications were identified in the examination, which help to illuminate this bias. In summary, this study discusses the importance of critically analyzing the tournament selection process due to the financial implications of an at-large tournament invitation. The results of this investigation provide vital information that could be utilized by the NCAA to monitor and adjust the current selection process.

The National Collegiate Athletic Association (NCAA) Division I Men's Basketball Tournament or "Big Dance" is one of the nation's largest sporting events. Driven by fervent university and regional affiliations, the tournament is followed intensely by fans across North America. For example, the 2008 final game between the University of Memphis and Kansas University was watched by over 19.5 million viewers (Seidman, 2008). This is an exceptionally large audience when compared to the decisive games of the National Basketball Association Finals (17.5 million) and National Hockey League Finals (6.8 million) (Brown, 2008; Street & Smith's SportsBusiness Daily [SBD], 2008b). In addition, the Alamo Dome in San Antonio, Texas witnessed unprecedented attendance figures at the 2008 Final Four including over 167,000 unique visitors to the Big Dance Free Street Party (SBD, 2008a).

Along with the intense media coverage and enormous fan following comes several significant implications for individual NCAA teams and athletic conferences. The most important of which is the financial benefits associated with making the field. The payouts for tournament participation were approximately $206,000 per round appearance in 2008. This sum is paid to the participating team's conference every year for six years, making the total amount approximately $1.24 million (NCAA, 2008a). More importantly, of the nearly $360 million distributed to individual conferences in 2008, nearly 40% ($143 million) of the total came from the Basketball Fund which distributes revenue to Division I conferences and individual institutions based on performance in the Division I Men's Basketball Championship (NCAA, 2008b).

Selection into the tournament also enhances recruiting opportunities through prolonged program visibility and provides extremely valuable public relations for a university. For example, the 2008 tournament performance by Davidson College resulted in the following phenomena: (1) the average daily sales at Davidson College Bookstore prior to Sunday, March 23, 2008 were approximately $1,700. Daily sales at Davidson College Bookstore on Wednesday, March 26, 2008 (the first day "Sweet 16" t-shirts were available) totaled $35,000; (2) the percentage increase in transfer inquiries received by Davidson's Admission Office since their second round win over Georgetown has been over 1,200 percent, and finally, (3) during the month of March 2008, Davidson College saw traffic on their website increase 262 percent (Davidson College, 2008).

Due to the significant benefits associated with tournament participation, interest in the selection of the tournament field is extraordinarily high. Consequently, the NCAA Division I Men's Basketball Tournament selection process has become a highly scrutinized system. The current process grants 31 teams an automatic berth into the tournament by winning their respective conference tournaments or by winning their conference regular-season title. A selection committee then decides the remaining 34 at-large bids. An at-large selection to the NCAA men's basketball tournament is defined as a bid or berth into the tournament granted by invitation. Currently, the NCAA follows a 25-step process for selecting the 34 best at-large bids (NCAA, 2008c). This process does not establish guidelines for evaluation or selections. It is a procedural tool used to ensure the selection committee is consistent from year to year.

There are many factors involved in selection decisions (including the Sagarin and RPI rankings), and specific information about the importance of certain attributes is not made public (Jing & Cox, 2006). Furthermore, the NCAA (2008c) holds that "the tournament bracket each year is based on the subjectivity of each individual committee member to select the best at-large teams available and to create a nationally balanced championship" (p. 8). As a result, critics of this process believe that there far are too many subjective biases involved given the magnitude of these selections.

Therefore, it is important to examine the committee's behavior in previous selections in order to understand the attributes that influence at-large selection. The purpose of this study was to examine the factors that influence at-large selection into the NCAA Division I Men's Basketball Tournament over a nine season period (1999-2007) in order to better to understand selection decisions. Both direct variables (i.e., team success and schedule attributes) and indirect variables (i.e., geographic location, team history, and local population) were analyzed to identify potential selection biases. These results will help determine whether selections are aligned with the criteria developed by the NCAA, and assist teams in improving their "resume" for future seasons.

The remaining sections of the paper are organized in the following manner. First, previous studies centering on the NCAA Division I Men's Basketball Tournament are discussed. Next, the composition of the selection committee, the criteria used to select at-large teams, and the potential biases that may result are established. Third, the methodology to test for empirical evidence of biased selection practices by the NCAA selection committee is provided. Finally, the results and implications of the analysis are discussed.

Review of Literature

Due to the enormous popularity of college athletics and the alleged improprieties of a flawed governing system, the examination of NCAA policy is hardly a new topic. Often at the focal point of this scrutiny is the lucrative NCAA Division I Men's Basketball Tournament. There have been several academic studies centering on the tournament from economic analyses of cartel behavior (DeSchriver & Stotlar, 1996) to gambling concerns (Colquitt, Godwin, & Caudill, 2001). In addition, numerous studies have used complex econometrics to predict margins of victory and final four probabilities of teams already selected (Carlin, 1996; Caudill, 2002; Schwertman, McCready, & Howard, 1991; Schwertman, Schenk, & Holbrook, 1996; Smith & Schwertman, 1999).

However, in terms of seeding and selection process issues, only a few studies have been performed. Fanning, Pilkington, Conrad, and Sommers (2003) assessed the seeding accuracy of the selection committee, while Harville (2003) provided a least squares system to solve for seeding obtrusiveness. In 2001, Coleman and Lynch examined the selection committee's accuracy in identifying at-large bids from 1994 to 1999. The researchers found six statistically significant variables (RPI rank, conference RPI rank, top 25 wins, conference wins-losses, top 50 wins-losses, & top 100 wins-losses) that affect the selection process and investigated yearly misclassifications. While comprehensive in obtaining direct variables, Coleman and Lynch did not examine any indirect factors such as geographic location, history of tournament appearances,

or local population. Given the subjective component of the committee's selection decisions, the examination of indirect variables may provide evidence of a need for refinement of the current method for at-large selection.

Jing and Cox (2006) extended the findings of Coleman and Lynch (2001) through the proposal of a machine learning approach to forecasting the behavior of the selection committee. Specifically, the researchers attempted to solve the seeding and at-large selection problems that stem from the NCAA's imperfect method of forming the field of 65 teams. The findings showed that the computer-based methods developed by the authors were competitive with the human predictors of the CollegeRPI.com and Sagarin rankings using data from 2005. The results also suggested that the committee based its decisions on a small number of factors, most notably the ranking features of Sagarin and the RPI. However, this study was limited to one year's worth data; therefore, to avoid bias, the authors recommended that future studies should include a larger data set.

Selection Committee

The NCAA Division I Men's Basketball Championship is under the control, direction, and supervision of the NCAA Division I Men's Basketball Committee. According to the NCAA (2008c), each of the 10 committee members serves a five-year term with two members rotating out and two new members rotating in each year. Committee representation also attempts to keep a geographical balance, as there are no fewer than two members representing the East, Midwest, South, and West regions at any one time. Individual institutions nominate qualified athletic administrators through their conference office. The NCAA Board of Directors appoints administrators from the list of nominees to serve on the selection committee for a given term (NCAA, 2008d).

One of the primary principles and practices of the committee is to "ensure that fair and equitable criteria are used to select the most deserving at-large teams," while also focusing on "administering a fair and equitable tournament by creating a nationally-balanced bracket comprised of the most deserving at-large teams and automatic-qualifiers chosen by conferences, while assigning institutions to sites as near to their campuses as possible" (NCAA, 2008c, p. 9). The selection committee is also responsible for maintaining the fiscal integrity of the tournament.

Selection Criteria

As mentioned above, the committee rarely provides specific information regarding selection of individual teams and releases only the final bracket to the public. However, specific direct factors certainly receive more attention from the fans and media as being central to the committee's selections. For instance, the RPI has garnered the majority of publicity as the main determinant for which teams are chosen. The RPI was created by the NCAA in 1981 to objectively assess men's and women's college basketball teams. The formula for this comprehensive ranking is 25 percent team winning percentage, 50 percent opponents' average winning percentage, and 25 percent opponents' opponents' winning percentage (Pomeroy, 2008). In 2004, the RPI was updated to account for location of wins. As a result, a home win now

counts as 0.6 of a win, while a road win counts as 1.4 wins. Conversely, a home loss equals 1.4 losses, while a road loss counts as 0.6 of a loss. The RPI currently ranks all NCAA Division I basketball programs in descending order from one to 343 based on their individual RPI score, so the best teams according to the formula are ranked closer to one, but have a larger RPI score.

In 2005, Greg Shaheen, the NCAA's vice president of Division I men's basketball, divulged the following general tools that are used by committee members to evaluate performance for selection purposes (in no particular order): win-loss record, overall RPI, non-conference record, non-conference RPI, conference record, conference RPI, road record, record in last 10 games, record against teams sorted by RPI, quality wins, quality losses, bad wins, bad losses, strength of schedule, and any other circumstances that could have affected results (Winn, 2005). While many of these attributes are easily evaluated and relate directly to on-court performance and/or administrative decisions, the final element (any other circumstances that could have affected results) opens the door for critics to assume that indirect variables or subjective biases may play a role in the selection of at-large teams. There is a human element to the selection process that allows for committee members to take personal observations into account. While the committee attempts to curtail personal biases by requiring committee members to excuse themselves from the room when teams directly-associated (institutional or conference) are being discussed, the system is not perfect. Given the subjective nature of this process, the following section highlights several potential biases that have recently received national attention and provides a rationale for the current study's selection of indirect variables.

Potential Selection Biases and Indirect Variables

It is difficult to objectively compare teams from conferences with varying levels of competitiveness. Given the competitive imbalance that dominates contemporary NCAA competition (Dennie, 2007), the fact that a majority of at-large bids consistently go to teams in major conferences has become a controversial topic. For instance, for the second consecutive season, only six of the 34 at-large teams selected in 2008 came from outside college basketball's six major conferences. By definition, the six "major" conferences in college basketball are the same six conferences that receive Bowl Championship Series bids in college football. They are the Atlantic Coast Conference (ACC), the Big 12 Conference, the Big East Conference, the Big Ten Conference, the Pacific 10 Conference, and the Southeastern Conference (SEC).

According to Thomas J. O'Connor (2007-2008 NCAA Selection Committee Chair), the committee looks at each team's individual resume and attempts to ignore conference affiliation. In addition, O'Connor noted that the committee does not focus on the number of bids by a certain conference. However, some teams and conferences with reputable credentials have been notoriously overlooked or "snubbed" by the selection committee. For example, teams representing the Missouri Valley Conference (MVC), more than any other conference, have fallen victim to the committee's imperfect practices. The MVC has had five RPI top 40 teams left out of the tournament. Missouri State University, particularly, has been left out of the tournament despite remarkable RPI's of 36, 34, and 21 (the only RPI top 25 team to not receive an at-large bid in the tournament's history).

Researchers Zimmer and Kuethe (2008) examined the potential for major conference bias in the seeding process of the NCAA's selection committee and determined that for the most part the selection committee does an adequate job of correctly seeding the NCAA Men's Division I Basketball Championship. However, the authors did suggest that "there is evidence of bias with respect to whether or not a team is a member of a power conference or a smaller conference" (p. 4). In an interview discussing how the committee handles a team with a stellar record from a weak conference, a former committee member stated, "I personally think this is one of the hardest things the committee has to do. You've got to project how a team may or may not compete if they were in a different conference" (Katz, 2008, ¶ 14). This is an excellent example of the potential for committee member bias. The projection of future performance is a subjective estimate and not an exact science. Therefore, this study investigated conference classification as an indirect variable that has a potential influence on the selection process.

Institutional location is another potential bias for committee members. Given the East Coast location of several highly-influential college sports media providers such as ESPN and CBS, and the residential location of several prominent national pollsters, the notion of a bias towards teams located in the East has received a lot of attention. While academic literature in this area is non-existent, columnists and fans debate this topic weekly through newspaper columns, Internet sites, and media blogs. The idea that games occurring in the Pacific Time Zone end at a time when most people on the East Coast are fast asleep is often the focal point of this argument. This time discrepancy often diminishes the amount of coverage awarded to teams on the West Coast, and ultimately, affects the public perception of the team. Given this intriguing issue, the current study investigated the relevance of geographic location as a potential bias.

In addition to examining the geographical region as an indirect variable, associated predictors were added to the model to investigate properties that may result in potential selection bias. As mentioned above, the selection committee is responsible for the fiscal viability of the tournament, and attracting and maintaining large audiences at tournament sites throughout the event are important means of creating revenue. In addition, the purchase power of these attendees is a vital component in the amount of money raised at tournament locations. Therefore, county population and per capita income variables were studied due to the potential economic significance of selecting institutions with a highly-populated and affluent local population.

Lastly, the landscape of Division I men's basketball is made up of traditional powerhouse teams. These teams provide a large fan base and national popularity for the tournament each year. As a result, there is a potential for recent tournament success of an at-large team to influence the selection committee's decision. However, according to former committee chair Craig Littlepage, past performances have no bearing on the brackets (Associated Press, 2008). Regardless, the current study examined a team's historical record of tournament appearances as a potential indirect variable in order to identify any influence that recent success has on at-large selection into the tournament.

In summary, the academic literature regarding at-large selection into the NCAA Division I Men's Basketball Championship is underdeveloped. Examinations of the direct variables associated with tournament selection are limited, and the analysis of indirect variables and subjective bias is non-existent. Due to the increasing financial benefits of tournament selection, it

is important to investigate the influence of both direct and indirect variables, as well as the fairness of the overall selection process.

Methodology

At-Large Selection Model

The current at-large selection model was developed to understand the effects that specific team success, selection criteria, and demographic variables have on selection into the NCAA Division I Men's Basketball Championship. This model was intended to provide empirical evidence regarding the impact that the factors being examined may have on tournament selection. Previous literature in the area of NCAA tournament selection (Coleman & Lynch, 2001; Jing & Cox, 2006) identified 19 team success and selection variables that may influence at-large selection. However, the current model is unique in comparison to previous models in that variables were included to identify the presence of any indirect influence in the selection process. Indirect variables were defined as variables that are not included in the list of tools used by the committee to evaluate team performance, but may still have a circuitous effect on selection into the tournament. The following section defines the variable of interest, at-large tournament selection, and the 17 explanatory variables examined in the current study. The explanatory variables are broken down by direct and indirect influence on tournament selection.

Dependent Variable

At-Large Selection (BID) – A dichotomous variable that identifies whether a team was given an at-large bid to the NCAA Division I Men's Basketball Championship in a specific year.

Direct Explanatory Variables

1. Winning Percentage (WINP) – The winning percentage for a team in a given year (Wins/Total Games Played).
2. Ratings Percentage Index Ranking (RPIRANK) – The final regular season RPI ranking for a team in a given year.
3. Pre-Season Ranking (PSRANK) – The Associated Press (AP) ranking that each team had prior to the beginning of a given season.
4. Strength of Schedule (SOS) – A value determined by the NCAA within the RPI that measures the difficulty of opponents for a team in a given season.
5. Winning Percentage during the Last Ten Games (WPLTG) – The winning percentage for a team during their last ten games of a given season. This variable includes conference tournament games in addition to regular season games.
6. Road Winning Percentage (WINPA) – The winning percentage for a team away from home during a given season.
7. Quality Wins (QUALWIN) – The number of wins for a team against opponents in the RPI top 50 during a given season.
8. Significant Losses (SIGLOS) – The number of losses for a team against opponents with an RPI ranking between 50 and 100.

9. Bad Losses (BADLOS) – Losses against teams with an RPI ranking of 100 or below.
10. Conference Regular Season Finish (CONFRSF) – A variable that identifies where a team finished in the regular season standings in their conference. Due to the large number of teams within some conferences, regular season finish was broken down into three categories: regular season champion, regular season top finish (1 to 4), and regular season bottom finish (5 or below).
11. Conference Tournament Finish (CONFTF) – A variable that identifies where a team finished in the conference tournament during a given season.

Indirect Explanatory Variables

12. Conference Classification (CONF) – The selection committee has stated that conference affiliation is not a factor in the selection process (i.e., each team's resume stands alone). However, affiliation with a conference that is traditionally successful may have an indirect influence on selection. Due to the large number of Division I conferences; the conference variable was broken down into three classifications: major (BCS conferences), mid-major (non-BCS conferences that had at least one at-large bid during the years examined in the current study), and small (non-BCS conferences that received zero at-large bids during the years examined in the current study). Table 1 provides a summary of conference classifications.
13. Geographic Region (GEO) – The region, divided by time zone, that a specific team is located. Certain geographic locations may be favored in the selection process.
14. Income (INCOME) - Given the importance of attracting affluent fans to tournament locations, this variable measured the per capita income of the county in which the institution is located in order to measure the potential purchasing power of the surrounding area.
15. Population (POP) - Attracting a large audience is vital for tournament organizers. Therefore, the population, as reported by the U.S. Census Bureau, of the county where a team plays was studied to investigate a potential population bias.
16. Tournament Appearance in the Previous Year (TAPY) – A dichotomous variable that identifies whether a team appeared in the tournament the year previous to a given year. The purpose of this variable was to identify any influence that previous selection has on current selection.
17. Tournament Appearance during the Previous Ten Years (TAPTY) - A dichotomous variable that identifies whether a team appeared in the tournament in the last ten years prior to a given year. The purpose of this variable is similar to that of TAPY, but it is a measure of historical tournament selection success.

Sample

The sample for the current study consisted of all NCAA Division I men's basketball teams that were ranked within the top 100 in the final RPI from 1999-2007. Teams in the final RPI top 100 that received automatic bids during this time period were excluded because they could not be selected for an at-large bid. All remaining teams during each year of the analysis were included in the data set. A total of 173 teams were represented in the sample. However, multiple teams finished in the final RPI top 100 more than once during the years being examined. A school's candidacy for the tournament was unique every year and therefore each year was counted as a separate observation. For example, one "observation" was Oregon's 2002 team that finished 23-8 with a final RPI ranking of 32, and received an at-large bid. A second "observation" was Oregon's 2004 team that finished 15-12 with a final RPI ranking of 89, and

did not receive an at-large bid. A total of 695 observations ($N = 695$) were included in the analysis. This sample provided a diverse set of teams from major, mid-major, and small conferences.

Table 1 – *List of NCAA Division I Basketball Conferences by Classification*

Major Conference	*Mid-major Conference*	*Small Conference*
Atlantic Coast Conference	Atlantic 10 Conference	America East Conference
Big 12 Conference	Colonial Athletic Association	Atlantic Sun Conference
Big East Conference	Conference USA	Big Sky Conference
Big Ten Conference	Horizon League	Big South Conference
Pacific-10 Conference	Mid-American Conference	Big West Conference
Southeastern Conference	Missouri Valley Conference	Ivy League
	Mountain West Conference	Metro Atlantic Athletic Conference
	West Coast Conference	Mid-Eastern Athletic Conference
	Western Athletic Conference	Northeast Conference
		Ohio Valley Conference
		Patriot League
		Southern Conference
		Southland Conference
		Southwestern Athletic Conference
		Summit League
		Sun Belt Conference

Procedure

Data on the variables in the at-large selection model were collected from a variety of sources. Team success variables such as winning percentage, RPI ranking, pre-season ranking, strength of schedule, regular season and conference tournament finish, road record, quality wins, significant losses, bad losses, and record in the last ten games were collected from ESPN.com and NCAA.com. Selection variables such as current at-large selection, tournament selection in the previous year, tournament selection in the previous ten years, and conference classification were collected from ESPN.com. Finally, demographic variables such as per capita income, county population, and geographic location were collected from the U.S. Census Bureau and Time.gov.

Statistical Design

A logistic regression was chosen because it is the most appropriate regression technique when the dependent variable is dichotomous (Hosmer & Lemeshow, 2000). For the current examination a direct (standard) logistic regression was conducted. The direct logistic regression technique allows all predictors in the evaluation to be entered simultaneously. According to

Tabachnick and Fidell (2007), this procedure should be used when there are no specific hypotheses about the order or importance of predictors. Due to the exploratory nature of this selection model, the direct logistic regression was the most appropriate statistical technique.

Prior to regression analysis, preliminary assessments were conducted on the variables in the model. Descriptive statistics and correlations were evaluated to identify any assumption or multicollinearity issues. Descriptive statistics and residual plots of the predictor variables showed no standard violations of regression which indicated sufficient power for the model. In addition, the dependent variable was binary, and the continuous predictors in the model were linear with the logit form of the dependent variable. Finally, the ratio of observations to variables was approximately 40 to 1 which was sufficient for this statistical technique (Tabachnick & Fidell, 2007). Potential multicollinearity issues within the model were also examined by investigating variance inflation factors and tolerance statistics. The results suggested there were no multicollinearity issues in the final regression equation used in the analysis.

Four areas of the regression results were evaluated. First, a goodness of fit chi-square (X^2) was used to test the significance of the overall model. In addition, the Nagelkerke index was evaluated for goodness of fit. This test is similar to that of r^2 in linear regression, but should not be looked at as explained variance in the dependent variable (Cohen, Cohen, West, & Aiken, 2003). However, a value closer to 1.0 is a measure of a satisfactory goodness of fit. Second, the Wald statistic and odds ratios were evaluated to understand significance and importance among the individual predictors in the model. The Wald test provided a test of significance for each predictor. The odds ratios provided information on the odds that there will be a change in category within the outcome variable if the independent variable increases by one unit (Tabachnick & Fidell, 2007). For example, if the odds ratio for quality wins were 1.5, it would indicate that for every additional quality win, the odds of being selected for the tournament are 1.5 times greater. Finally, misclassifications were examined through the evaluation of predicted versus expected outcomes. A significance level of 0.05 was established *a priori* in analyzing the regression model and related variable correlations.

Results

Descriptive statistics were analyzed to compare characteristics of teams in the data set that received at-large tournament bids to teams that did not. The results showed that team success variables generally favored the teams selected for an at-large bid into the tournament. Selected teams had a .723 winning percentage compared to .606 for non-selected teams. Selected teams also had more quality wins (5.12) and less significant losses (1.08) compared to non-selected teams (2.16 and 2.36, respectively). However, strength of schedule differed only slightly (.562 for selected teams compared to .538 for non-selected teams). In addition, selected teams had more historical success in terms of tournament selection (5.17 tournament appearances in the last 10 years for selected teams compared to 2.95 for non-selected teams). Table 2 provides a complete list of descriptive statistics on the continuous predictor variables in the model broken down by tournament selection.

Table 2 - *Descriptive Statistics*

Variable	Variable Description	Mean	Standard Deviation	Minimum	Maximum
Selected Teams					
WINP	Winning %	.723	.0781	.533	.964
RPIRANK	RPI Ranking	26.84	14.959	1	74
PSRANK	Pre-Season Ranking	19.68	12.962	1	58
SOS	Stength of Schedule	.5619	.0295	.4766	.8180
TAPTY	Tournament Appearances in the last 10 years	5.17	3.037	0	10
WPLTG	Winning % in Last 10 Games	.655	.1386	.300	1.000
WINPA	Winning % Away from Home	.566	.1284	.231	.929
QUALWIN	RPI Top 50 Wins	5.12	2.480	0	14
SIGLOS	RPI 50-100 Losses	1.08	1.278	0	13
BADLOS	RPI 100 and Below Losses	.23	.6330	0	5
INCOME	Per Capita Income	21658	4346.39	13421	21659
POP	County Population	916223	1839697	7422	9948081
Non-Selected Teams					
WINP	Winning %	.606	.0882	.357	.897
RPIRANK	RPI Ranking	73.46	17.493	21	100
PSRANK	Pre-Season Ranking	32.47	13.863	6	59
SOS	Stength of Schedule	.5382	.0311	.4531	.6039
TAPTY	Tourny Appearances in the last 10 years	2.95	2.363	0	10
WPLTG	Winning % in Last 10 Games	.543	.1714	.100	.900
WINPA	Winning % Away from Home	.429	.1494	.000	.857
QUALWIN	RPI Top 50 Wins	2.16	1.620	0	8
SIGLOS	RPI 50-100 Losses	2.36	1.641	0	9
BADLOS	RPI 100 and Below Losses	.70	1.070	0	6
INCOME	Per Capita Income	21628	4687.65	13421	42922
POP	County Population	863976	1419885	5620	9948081

Note: *N* = 695

The 17 predictors were made up of a combination of continuous and categorical variables. Correlations were examined for each of the 12 continuous predictor variables and at-large selection into the tournament. Ten of the twelve continuous independent variables were significantly correlated to at-large selection at the 0.05 level. The continuous independent variables that had the strongest correlation with at-large selection were RPI ranking ($r = -.816$), quality wins ($r = .584$), winning percentage ($r = .572$), pre-season ranking ($r = -.416$), and significant losses ($r = -.390$). In addition, chi-square analyses were conducted for each of the five categorical predictor variables and at-large selection. Three of the five categorical variables (conference classification, tournament appearances in the previous year, and conference regular season finish) had a significant relationship with at-large selection.

Direct Logistic Regression

The results of the direct logistic regression analysis indicated that the at-large selection model was found to be significant ($X^2(23) = 760.34$, $p < .001$). The classification table assessing fit of the model showed that 93.5% of observations were correctly classified. In addition, the model had a Nagelkerke r^2 score of .891, indicating a satisfactory goodness of fit.

Table 3 shows Wald statistics and p-values for each of the 17 predictors. According to the Wald criterion, there were five significant variables within the model: RPI Rank (X^2 [1, 695] = 34.52, $p < .001$), conference classification (X^2 [2, 695] = 26.06, $p < .001$), quality wins (X^2 [1, 695] = 16.81, $p < .001$), conference regular season finish (X^2 [2, 695] = 9.18, $p = .010$), and significant losses (X^2 [1, 695] = 5.55, $p = .019$). The majority of these predictors are direct variables that the selection committee acknowledges as evaluative tools for selection. Conference classification was the only significant variable that may have an indirect influence on tournament selection.

Odds ratios and confidence intervals for the five significant predictor variables were subsequently examined. The odds ratios for quality wins, significant losses, and RPI rank were noteworthy. Quality wins had an odds ratio of 1.77, which indicated that a team is approximately 77% more likely to be given an at-large invitation for every additional win against an opponent in the RPI top 50. Significant losses also had a considerable odds ratio of .652, which indicated that a team is 35% less likely to be given an at-large invitation for every additional loss against an opponent ranked 50 to 100 in the RPI. Finally, RPI rank had an odds ratio of .854 which indicated that a team is 15% less likely to receive an at-large bid for every one-unit increase in RPI ranking (i.e., a move from 24 to 25 in the RPI rankings).

Finally, misclassifications were analyzed to identify specific teams that, based on the model, were not categorized accurately. An analysis of predicted versus observed cases was conducted. Table 4 provides a summary of all misclassifications in the sample. The results showed that there were a total of 27 misclassifications from 1999 to 2007. There were at least two misclassifications in all years of the study except 2002. The 2006 tournament had the most misclassifications (5), followed by 2000 and 2007 (4 each). There were a total of fifteen misclassifications involving teams that were identified as mid-major. There were also twelve major conference misclassifications. There were no misclassifications among small conference teams.

Table 3 - *Direct Logistic Results (Dependent Variable = At-Large Tournament Selection)*

Variable	Wald Chi-Square statistic	p-value
WINP	1.819	.177
RPIRANK	34.516	<.001
PSRANK	1.140	.993
SOS	.952	.329
TAPY	.061	.804
TAPTY	.985	.321
WPLTG	2.124	.125
WINPA	1.927	.165
QUALWIN	16.812	<.001
SIGLOS	5.547	.019
BADLOS	2.104	.147
CONF	26.057	<.001
CONFRSF	9.182	.010
CONFTF	3.667	.300
INCOME	1.983	.159
POP	.770	.380
GEO	3.203	.361

Table 4 – *List of Misclassifications 1999-2007*

Year	Team	Final RPI	Status
1999	Univ. of Mississippi	51	Selected
1999	Univ. of Nebraska	47	Not Selected
1999	Univ. of New Mexico	74	Selected
2000	Indiana State Univ.	69	Selected
2000	Southern Methodist Univ.	49	Not Selected
2000	SW Missouri State Univ.	34	Not Selected
2000	Vanderbilt Univ.	39	Not Selected
2001	Univ. of Missouri	47	Selected
2001	Univ. of Southern Mississippi	53	Not Selected
2003	Auburn Univ.	36	Selected
2003	NC State Univ.	53	Selected
2003	Boston College	49	Not Selected
2004	Univ. of Richmond	47	Selected
2004	Colorado Univ.	57	Not Selected
2005	NC State Univ.	65	Selected
2005	Univ. of Northern Iowa	37	Selected

2005	Oral Roberts	50	Not Selected
2005	Univ. of Alabama-Birmingham	49	Selected
2006	Air Force	50	Selected
2006	Creighton Univ.	37	Not Selected
2006	Hofstra Univ.	30	Not Selected
2006	Missouri State Univ.	21	Not Selected
2006	Seton Hall Univ.	58	Selected
2007	Air Force	30	Not Selected
2007	Old Dominion Univ.	39	Selected
2007	Stanford Univ.	65	Selected
2007	Syracuse Univ.	50	Not Selected

Discussion

The primary purpose of this study was to examine the factors that influence at-large selection into the NCAA Division I Men's Basketball Championship. Previous research has focused on variables that have a direct influence on selection (Coleman & Lynch, 2001; Jing & Cox, 2006). However, these examinations have not considered potential indirect variables that may also impact these decisions. The current study analyzed both types of variables over a nine year period (1999-2007).

The results suggested that, for the most part, selections matched the performance-based guidelines used by the selection committee. The variables RPI rank, conference classification, conference regular season finish, quality wins, and significant losses were found to be significant predictors of selection. All of these variables, with the exception of conference classification, were consistent with the statement made by Greg Shaheen, NCAA's vice president of Division I Men's Basketball and the findings of Coleman and Lynch (2001).

However, there were several issues which arose from this investigation. First, some of the other variables that were cited by Shaheen as important in the selection process, such as strength of schedule and wins in the last ten games, were not found to be significant predictors of tournament selection. This information was also consistent with the findings of Coleman and Lynch (2001). The results provide evidence that perhaps the committee tends to focus more attention on specific direct variables of interest.

Second, conference classification was found to have a significant influence on tournament selection. This may provide evidence of a major conference bias, which is supported by specific misclassifications over the nine-year period. The issue of a potential conference bias is substantial, considering the financial benefits and exposure associated with tournament selection. A closer look at the misclassifications illuminates this possible conference bias. Of the five teams that should have been selected for the NCAA tournament in the past two seasons (based on the current analysis), four are considered mid-majors: Creighton (RPI = 37), Hofstra (RPI = 30), Missouri State (RPI = 21), and Air Force (RPI = 30). These schools were left out in favor of major conference schools such as Seton Hall (RPI = 58) and Stanford (RPI = 65). One possible explanation for this bias could be that lower levels of media exposure during the regular

season for small and mid-major conference teams do not provide the national publicity that major conference teams enjoy. Another possible explanation is that only two NCAA champions in the current tournament format have come from mid-major conferences. This may influence selection due to the fact that teams are given at-large bids based, in part, on projected success in the tournament.

Although, the committee has been consistent in terms of selection criteria, the significance of misclassifications cannot be overstated. This is especially true when a major conference team is chosen over a mid-major conference team. It is important to understand the considerable financial benefits associated with tournament selection, specifically for mid-major or small conference teams. The direct financial benefit in terms of the payout from the NCAA's Basketball Fund is the first thing that these schools are denied. Each game appearance is worth well over one million dollars to that school's conference, and this money is typically shared among schools in that conference. With athletic department budgets that are considerably less than the major conference schools, this revenue is much more important for mid-major programs.

In addition, numerous studies have concentrated on the effects of intercollegiate athletic prowess on increased enrollment applications and increased donations to the university (Baade & Sundberg, 1996; Goff, 2000; Sigelman & Bookheimer, 1983; Sigelman & Carter, 1979; Sperber, 2000; Zimbalist, 1999). While the results are mixed on whether athletic success translates significantly into direct benefits for a university, perceptions still exist that recent on-court success can create indirect benefits for a university.

Consider George Mason University, the NCAA Tournament's 2006 Cinderella team. A closer look at their athletic department budget indicates that they generated over $8.4 million in student fees and $1.38 million in direct and indirect institutional support as their two primary sources of revenue in 2006 (total revenues amount to slightly over $12.1 million). According to the Commonwealth of Virginia Auditor of Public Accounts (2006), the George Mason University Athletics Department still lost over $570,000 in 2006. Had Hofstra and Old Dominion (both in the Colonial Athletic Association with George Mason) been given at-large invitations, George Mason would have received additional revenue to offset their losses. The institutional support they received could potentially have gone to other academic programs rather than drawing down athletic department debt.

The benefits of making the tournament are multiplied with each tournament victory. The NCAA's Basketball Fund pays teams for each game appearance, so a win equals another $206,000 a year for six years to that team's conference. However, the ancillary benefits are potentially even more profound. Within two years of George Mason's run to the Final Four, a recent study reported a 350% increase in admissions inquiries, a 500% increase in hits on the school's official website, a 52% increase in fundraising, a doubling of season ticket sales for the basketball team, and over $677 million in media coverage associated with the 2006 Tournament (Baker, 2008). Further, the school sold over $1 million in Final Four merchandise. Of course, not every mid-major team is going to see the success that George Mason University did in 2006; however, the potential for these benefits exist only if a team is selected for the tournament.

Based on these findings, it appears that although the tournament committee appears to be generally consistent with selection guidelines, the evidence of a potential major conference bias

suggests a need to refine the selection methods currently taking place. The implications of selection decisions have long-term financial effects for athletic departments, as well as the academic institutions as a whole. Schools and conferences are losing millions of dollars in direct and indirect revenues based on these selections. Ultimately, the current model illustrates the importance of empirical evaluation of a process that is subjective in nature due to the use of a selection committee. Therefore, it would benefit the committee to refine the current system of evaluation in order to limit the impact of indirect variables or subjective biases in the tournament selection process. Investigations such as the current study can help to illuminate potential bias and further enhance the accuracy of the selection process.

Suggestions for Future Research

The current model must be tested with different samples across varying time periods to increase its overall generalizability. Additionally, future research may help identify other variables not included in this study that further explain tournament selection. Additional direct and indirect variables should be considered in future analyses. Due to the human component found in the selection process, it is important to constantly examine selection methods in order to identify potential bias and refine the evaluation system regularly. Finally, while the literature on the benefits of tournament success is considerable, future research could examine the distribution and use of the Basketball Fund payouts to further quantify the impact of tournament success on conferences and their member universities.

References

Associated Press (2008). *Littlepage goes overtime to defend panel's work.* Retrieved November 8, 2007 from http://sports.espn.go.com/ncb/ncaatourney06/news/story?id=2367339

Baade, R.A., & Sundberg, J.O. (1996). Fourth down and gold to go? Assessing the link between athletics and alumni giving. *Social Science Quarterly, 77*, 789-803.

Baker, R. (2008, March 24). A funny thing happened on George Mason's way to Final Four. *Street & Smith's Sports Business Journal, 9*, 28.

Brown, M. (2008, June 6). *NBC reports 6.8 million viewers for game 6: Up 111% from 2006.* Retrieved on August 28, 2008, from http://www.bizofhockey.com/index. php?option=com_content&view=article&id=147:nbc-reports-68-million-viewers-for-game-6-up-111-from-2006&catid=40:television&Itemid=59

Carlin, B. P. (1996). Improved NCAA basketball tournament modeling via point spread and team strength information. *The American Statistician, 50*, 39-43.

Caudill, S. B. (2002). Heterogeneous skewness in binary choice models: Predicting outcomes in the men's NCAA basketball tournament. *Journal of Applied Statistics, 29*(7), 991-1001.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Coleman, B. J., & Lynch, A. K. (2001). Identifying the NCAA tournament "Dance Card." *Interfaces, 31*(3), 76-86.

Colquitt, L. L., Godwin, N. H., & Caudill, S. B. (2001). Testing efficiency across markets: Evidence from the NCAA basketball betting market. *Journal of Business Finance & Accounting, 28*(1-2), 231-248.

Commonwealth of Virginia Auditor of Public Accounts (2006). *George Mason University intercollegiate athletics programs for the year ended June 30, 2006.* Retrieved October 1, 2008, from http://www.apa.state.va.us/reports/GMUncaa06.pdf

Davidson College. (2008). *By the numbers.* Retrieved on April 16, 2008, from http://www.davidson.edu/news/BytheNumbers.html

Dennie, C. (2007). White out full grant-in-aid: An antitrust action the NCAA cannot afford to lose. *Virginia Sports & Entertainment Law Journal, 7,* 97-125.

DeSchriver, T. D., & Stotlar, D. K. (1996). An economic analysis of cartel behavior within the NCAA. *Journal of Sport Management, 10*(4) 388-400.

Fanning, C. E., Pilkington, K. J., Conrad, T. A., and Sommers, P. M. (2003). March madness: The racket in regional brackets. *Journal of Recreational Mathematics*, *31*(3), 201-207.

Goff, B. (2000) Effects of university athletics on the university: A review and extension of empirical assessment. *Journal of Sport Management, 14,* 85-103.

Harville D. A. (2003). The Selection or seeding of college basketball or football teams for postseason competition. *Journal of the American Statistical Association, 98*(461), 17-27.

Hosmer, D.W., & Lemeshow, S. (2000). *Applied logistic regression.* New York, NY: John Wiley & Sons, Inc.

Jing, Y., & Cox, A. (2006). *Crashing the dance: Learning the behavior of the NCAA Basketball Selection Committee.* Retrieved on August 28, 2008, from http://www.cc.gatech.edu/~yjing/work/cs6230_project1.pdf

Katz, A. (2008). *Those in the selection know give a sneak peek.* Retrieved November 8, 2007 from http://sports.espn.go.com/ncb/columns/story?columnist=katz_andy&id=2338089.

NCAA. (2008a). *Current basketball revenue distribution information.* Retrieved on November, 14, 2008 from http://www.ncaa.org/wps/ncaa?ContentID=4

NCAA. (2008b). *Current Division I total revenue distribution information.* Retrieved on November, 14, 2008 from http://www.ncaa.org/wps/ncaa?ContentID=4

NCAA. (2008c). *NCAA division I men's basketball championship principles and procedures for establishing the bracket.* Retrieved on August, 27, 2008, from http://www.ncaa.com/uploadedFiles/Sports/Basketball_(M)/Articles/Division_I/NCAADivisionIMenPrinciplesandProcedures.pdf

NCAA. (2008d). *Division I Governance, Cabinet, Committee Vacancies.* Retrieved on February 1, 2009, from http://www.ncaa.org/wps/ncaa?ContentID=866

Pomeroy, K. (2008). *RPI help.* Retrieved on August 27, 2008, from http://kenpom.com/blog/index.php/weblog/rpi_help/

Schwertman, N. C., McCready, T. A., & Howard, L. (1991). Probability models for the NCAA regional basketball tournament. *The American Statistician, 45*, 35-38.

Schwertman, N. C., Schenk, K. L., and Holbrook, B. C. (1996). More probability models for the NCAA regional basketball tournaments. *The American Statistician, 50*, 34-38.

Seidman, R. (2008, April 8). *NCAA men's basketball final: Slightly down from last year...* Retrieved on April 16, 2008, from http://tvbythenumbers.com/2008/04/08/ ncaa-mens-basketball-final-slightly-down-from-last-year/3252

Smith, T., & Schwertman, N. C. (1999). Can the NCAA Basketball tournament seeding be used to predict margin of victory? *The American Statistician, 53*(2) 94-98.

Sigelman, L., & Carter, R. (1979). Win one for the giver? Alumni giving and big-time college sports. *Social Science Quarterly, 60,* 284-294.

Sigelman, L., & Bookheimer, S. (1983). Is it whether you win or lose? Monetary contributions to big-time college athletic programs. *Social Science Quarterly, 64,* 347-359.

Sperber, M. (2000). *Beer and circus.* New York: Henry Holt and Company.

Street & Smith's SportsBusiness Daily. (2008a). *San Antonio gets high marks, plans to bid for more Final Fours*. Retrieved on June 6, 2008, from http://www.sportsbusinessdaily.com/article/119903

Street & Smith's SportsBusiness Daily. (2008b). *ABC earns 10.1 fast national rating for NBA Finals game five.* Retrieved on August 28, 2008, from http://www.sportsbusinessdaily.com/article/121672

Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics (5th ed.).* Boston, MA: Pearson Education.

Winn, L. (2005, March 9). Inside 'the bunker': How the selection committee's work may unfold. *SI.com*. Retrieved on November, 14, 2008, from http://sportsillustrated.cnn.com/2005/ writers/luke_winn/03/09/inside.bkc/index.html

Zimbalist, A. (1999). *Unpaid professionals.* Princeton, NJ: Princeton University Press.

Zimmer, T., & Kuethe, T. (2008). Major conference bias and the NCAA Men's Basketball Tournament. *Economics Bulletin, 12*(17), 1-6.