# A Longitudinal Study of Regional Bracket Equality in the NCAA Men's Basketball Tournament

John A. Trono
Saint Michael's College
Colchester, VT
Tech report: SMC-2013-CS-001

## Abstract

Several articles have described certain mathematical/statistical models that were designed to accurately predict the teams that will be invited, as the non-automatic-qualifiers, to the NCAA men's basketball tournament. These are the positions in the tournament bracket that are not reserved for (designated) conference champions. There has also been recent research concerning which teams should have received those tournament invitations. These articles have proposed specific, objective methodologies for determining the teams that appear to be the most worthy recipients for such tournament participation. However, this paper will investigate several related topics, where the major focus will be: how evenly balanced have the four separate regions been in recent tournaments, in comparison to the previous methodology that was used to construct these tournament pairings? Unbeknownst to many, tournament game pairings and placement into a particular region, before 1979, were (mostly) predetermined before the season even began, so how evenly balanced was that tournament scheduling strategy when compared to current practices, i.e., were some geographic regions filled with more (fewer) strong teams back then than the other regions? Finally, the accuracy of the assigned tournament seeds, introduced in 1979, will also be examined.

After an extensive and comprehensive review of how the process in which the teams have been chosen to play in the NCAA tournament has evolved, and the expectations for how the tournament should unfold, the new metric developed here (the Tournament Selection Ratio), which may also be used to evaluate which teams deserve an invitation, will be presented. This will be followed by the description of: the model for computing the probability a team will reach subsequent rounds of the tournament; the application of this model to determine how fair the assignment of teams into specific regions has been, both before and after seeding became the common practice when aligning teams into the tournament brackets; and the concerns that some brackets (especially before 1979) were unfair to particular teams, regarding their path to the national championship. The results concerning how balanced the brackets have been over the years will conclude this report.

## Introduction

The men's National Collegiate Athletic Association (NCAA) basketball championship tournament was expanded from 8 to 16 teams in 1951. Of those 16 teams, 10 were awarded

automatic bids to that tournament, representing their specific conference as either the champion of the conference's post-season tournament, or having earned that bid because they possessed the best intra-conference, won-loss record that season. The 1951 NCAA tournament bracket appears to be the first one published before the regular season began, and the 10 conference champion slots were filled as just described. The appointed NCAA tournament committee selected the remaining six at-large teams "… on the basis of record, regardless of whether independent or a conference member" (Petersen, 1951). This tournament bracket template was referred to as the tournament draw back then, and a photocopy of one (from the *1951 Official NCAA Basketball Guide*) can be found at the end of this report. In 1975, the NCAA tournament was expanded to include 32 teams: 17 conference champions and 15 at-large teams. That was also the first year since 1952 (as will be supported shortly) where conference runner-ups could be invited to the tournament, and, where teams were not essentially assigned to specific regional brackets based primarily on the geographic location of their campus, or conference affiliation. This significant change in policy had an immediate impact as witnessed when two teams from the same conference (Indiana and Michigan) played in the 1976 championship game.

The men's NCAA basketball tournament expanded in 1979 to allow 40 teams to be invited; this increased to 48 teams in 1980, 52 in 1983, 53 in 1984, 64 teams in 1985, 65 in 2001, and finally 68 in 2011. Coincidentally, 1979 was also the first year when each team was assigned an official seed within its respective region. This strategy was (apparently) instituted as an effort to balance each bracket, in a manner similar to tennis tournaments, where the best players are not scheduled to play each other until the tournament's final rounds. The year 1979 has also sometimes been referred to as the break out year for this tournament. The amount of interest in that year's championship game, which pitted Michigan State, and their outstanding, sophomore point guard Earvin (Magic) Johnson, against the undefeated Indiana State team that was led by their senior, sharp shooting star Larry Bird, "… set TV ratings that still stand today" (Infoplease.com).

The decision to include more teams, thereby involving more fans directly (than those who typically followed this tournament's results, regardless of the specific teams invited), also provided increased incentive for expanded media coverage, i.e., nationally televising more games than just the regional finals, and the games between the Final Four. The 1982 championship game, featuring freshman shot-blocking sensation Patrick Ewing (of Georgetown), and coach Dean Smith's North Carolina team - that included future NBA greats Michael Jordan, James Worthy and Sam Perkins - helped to accelerate the fascination with this tournament. This high level of interest permeated the American world of sport in 1983 when Jim Valvano's 'Cinderella' North Carolina State team, only seeded sixth in their region, won the championship by defeating the explosive Houston Cougars (#1 in both polls) which included Hakeem Olajuwon and Clyde Drexler. (Both of the latter two championship games were close contests throughout, and the final outcomes were in doubt until the very last minute of each of those games, producing a compelling and captivating viewing experience.) Georgetown and Houston battled for the championship in 1984, pitting Patrick Ewing against Hakeem Olajuwon in a classic matchup of two great teams - and their seven foot centers. As a senior, Ewing led Georgetown back to center stage in 1985, his third championship game in his four collegiate years, only to lose to a #8 seed (Villanova) - that played an almost perfect game, as exemplified by their making 22 of their 28 (two point) field goal attempts - by a final score of 66 to 64.

All of those exciting tournaments finales, along with increasing media support, helped to accelerate this as a highly anticipated, yearly phenomenon so much so that many fans have now become completely obsessed with this month long sporting event: approximately five million people watched the 2005 'Selection Sunday' TV show, and an estimated 15 million tuned in to witness North Carolina play Illinois in the championship game that same year (Jing and Cox, 2008). After the tournament bracket is announced during the Selection Sunday show, the excitement truly gains significant momentum with the subsequent four day, basketball extravaganza where 48 games are played (most are at least partially televised) that following Thursday through Sunday - yielding the tournament's surviving Sweet Sixteen teams.

With this heightened level of media coverage, and more information being easily accessible over the Internet, which teams are awarded the at-large bids, and which are not, has been given much more scrutiny recently, especially regarding how successful recent NCAA selection committees have been when designating who they determined were the most deserving teams for those non-automatic qualifier tournament invitations. After reviewing the structure of previous NCAA tournaments (i.e., before 1979), the results regarding the determination if each region is essentially as balanced as the other three, will be tabulated for recent years as well as those predating the term 'March Madness'. Prior to the dissemination of these results, concerning the equality of the four regions (both past and present), will be the development of the respective tools and techniques that were applied to produce the conclusions related to the main topic here: bracket balance.

## History

The first National Invitational Tournament (NIT) was held in 1938, and the NCAA began its own single elimination tournament the following year, inviting four teams from each region (Eastern and Western), with the regional winners meeting for the championship somewhere other than where the two regional tournaments were held. This structure remained intact until the 1951 tournament, when the field was expanded to 16 teams, specifying 6 at-large teams to be determined along with the 10 conference affiliated, automatic bids. In 1952, the 'Final Two' was expanded to the first Final Four as the Eastern and Western regional championship games preceded the national championship game, and those last few games were all to be played at the same site; from 1952 through 1955, the two regions actually each had two separate brackets, where those winners met for said regional championship (at the Final Four site).

Unlike today, where both tournaments are essentially held concurrently, the NIT was completed prior to the NCAA tournament back then. Some teams were actually invited to play in both tournaments, and surprisingly, in the 1940's, some teams refused invitations to participate in one or both tournaments; that decision would be unthinkable today, given the monetary incentives from participating in the NCAA tournament. (The NIT and NCAA champions agreed to play against each other from 1943-1945. Those games were scheduled several nights after the NCAA tournament was finished, with the proceeds benefiting the Red Cross - as the end of World War II drew near.)

1952 was the last year where teams could compete in both post-season tournaments, as "… eligibility for the [NCAA] competition will be limited to members that compete in only one

postseason tournament" (Petersen, 1953). In 1953, the NCAA field was expanded from 16 to 22 teams. However, "To avoid the danger of cheapening conference championships, runner-up teams are not eligible for selection as at-large entries in the national playoffs." (Petersen, 1953) The four distinct sub-regions (East, Midwest, West, and Far West - the latter three being renamed Mideast, Midwest and West two years later) were established in 1956, though it remained that the Eastern regional champion played against the Western champion until 1973. The total number of invitations varied between 23 and 25 teams from 1954 until 1969, when it was set to 25 (including 9 or 10 at-large bids) - until the aforementioned jump to 32 teams in 1975 (and 15 non-conference champions). For 10 of those 16 years, the tournament draw had some built-in flexibility where: a tournament spot could either be filled by an at-large team, or a conference champion (Ivy in 1963 & 1964, Ohio Valley in 1956, or Yankee in 1967), depending upon the judgment of the NCAA selection committee as to the specific team's worthiness; or a place in the tournament could be left open, removing a first round game if it was deemed that enough quality, at-large teams were not available, in a particular geographical region. An at-large team was not selected for such a tournament bracket position in 1957, 1959, 1960 and 1968, though in 1958 a team was invited. Two such decisions were to be made in the 1961 tournament: one slot was filled and the other was not. The winners of the East and Mideast sub-regions were scheduled in the published tournament draw to play each other in the Eastern regional final during the Final Four (likewise with the Midwest and West winner in the Western regional final) until 1973, when the four sub-regions became four, truly distinct regions; that was also when a rotation began which established the regional pairings for each subsequent year's Final Four.

From 1954 to 1956, one sub-region in the Eastern region placed the Big Ten and Southeastern conference (SEC) champions into the second round, and each sub-region had between one and three teams that received a bye. The two games providing the opponents for the Big 10 and SEC champions typically included the Ohio Valley champion playing an at-large team, and likewise with the Mid-American conference (MAC) champion. If the game providing the Big 10 champion's opponent included the MAC champion, then the next year, that game's winner would play the SEC champion, so there was some rotation within these published tournament draws, to provide some non-biased, periodic variation. When the four sub-regions were officially named in 1956, this specific alignment, with the Big 10 and SEC champs, became the format for the Mideast sub-region's tournament draw. For comparison purposes, the East sub-region had one such bye, initially for the Ivy League's champion, then eventually this spot went to the representative from the Mid-Atlantic conference, and finally, in 1963, the Atlantic Coast Conference (ACC) champion assumed that position in the East's bracket. The order of the three other conference champions, that played at-large teams in the three first round games in the East's sub-region, rotated in a reasonably predictable manner, and similar patterns existed in the Midwest and West sub-regions as well. (The term region will now be synonymous with sub-region for the rest of this article.)

A large number of teams did not belong to any particular conference during these early years of the NCAA, and so, the at-large selections came from a large group of 'independents' - teams that were not affiliated with any conference. However, all four regions each included quite a disparate number of independent teams, which made the competition for those limited, at-large bids much fiercer in some regions than the others. The geographical boundary for the at-large bids in the Mideast and Midwest regions was unspecified, but teams tended to be chosen to play

in one region over the other. Notre Dame always played in the Mideast region, except in 1965 and 1971, when they competed in the Midwest region. Similar exceptions were made for Loyola of Illinois in 1968, for Dayton in 1969 and 1970, and for Marquette in 1961 and 1977 (more details concerning the 1977 tournament will appear later in this article). DePaul was invited into the Midwest region in 1959 and 1960, but competed in the Mideast region in 1966. Houston was also invited into the Midwest region many times in the late 1960s, and early 1970s, but they competed in the West region in 1966 as well as Colorado State that year, though Colorado State was invited into the Midwest region in 1967 and 1969. The opposite was true for New Mexico, which played in the Midwest in 1970, but was invited into the West region in 1968 and 1974; New Mexico State was invited into the West twice (1967 and 1971) and into the Midwest twice (1968 and 1969). Oklahoma City competed in the West region in 1965, even though they had been invited to the Midwest region in 1956, 1957, 1963, 1964 and 1966. Omitting any specific definitions for these geographic boundaries provided some selection flexibility for the committee, but it also invited the opportunity for some bias to (possibly) creep into these important decisions regarding the at-large slots. (Another example of this flexibility occurred in 1959. The team from the University of Portland - in Oregon - was invited to compete in the Midwest region, though it seems obvious that geographically, they should have been included in the West region.)

Perhaps in light of the anticipated expansion to 32 teams in 1975, South Carolina was invited into the Midwest region in 1973, having recently left the ACC, and two other teams travelled to play outside their respective, corresponding geographical regions in 1974 (though South Carolina was invited back into the East region in 1974): Dayton was invited to play in the West region, while Syracuse was invited to play in the Midwest region. This practice would not be considered unusual now since the present day tournament selection committee assigns teams into a region in an effort to make them evenly balanced, but the rules and practices from 1953 to 1974 did not normally allow for that to occur. When the number of invited teams was increased from 40 in 1979 (when the advent of seeding made the tournament draw idea obsolete) to 48 teams in 1980, the five year cap - limiting each conference to a maximum of two participating teams - was eliminated, and 6 conferences filled 22 of the 48 tournament invitations: 5 teams were from the ACC, 4 from the Big 10 and the Pac-10 conferences respectively, 3 more were from a much smaller Big East conference (7 teams in 1980 instead of the 16 present in 2010-11), 3 came from the SEC, and 3 from the Metro conference, which eventually became Conference USA. This dramatic shift, concerning which teams could be invited to compete in the NCAA Men’s basketball tournament, raises the questions: how well did each committee (both before and after 1979) evenly spread the talent inside each region each year, and, were some deserving teams left out, i.e., were the best remaining teams selected for the at-large bids once geographic location, and/or conference caps, were taken out of the process that informed such decisions?

## Desired Tournament Conclusion

As previously mentioned, tennis tournaments have traditionally seeded players in such a manner so that if upsets do not occur, then the players officially recognized as the best will meet in the finals (and semifinals). Players whose performance has earned them such recognition are rewarded by pairing them with opponents in earlier rounds of the tournament who have not demonstrated superior prowess that year, which increases the likelihood that said rewarded players will advance further in the tournament. Therefore, if everything goes according to plan,

the #1 seeded player faces the #2 seeded player in the finals, and the paying, courtside spectators will have the opportunity to watch this highly anticipated match. Of course, as we all know, upsets do occur, as do injuries both prior to (and during) some of these matches.

When the NCAA selection committee started seeding teams in a similar manner in 1979, using that season's collection of game results as evidence, the committee might have hoped that its four top teams (which they designated as the #1 seeds in each of the four regions) were going to reach the Final Four; however, it took almost 30 years for that to finally occur (in 2008), and it wasn't until 1993 that three #1 seeds actually earned their way there. Though they are considered to be somewhat subjective, before rating and ranking systems became officially utilized as useful tools to help guide the selection committee, the two polls - one representing the sportswriters who vote in the Associated Press (AP) poll, and the other, where certain coaches were invited to participate in the United Press International's (UPI) poll (this particular poll is now overseen by ESPN and the *USA Today*) - can also be considered as useful measures to evaluate how close the Final Four was to being (who those voters thought were) the top four teams in the country before the tournament began.

The AP poll began in January, 1949, and the UPI poll started two years later. A game featuring the #1 team versus the #2 team in the NCAA championship game actually did occur in 1949, but that special event didn't reoccur until 1957. That triple overtime thriller, won by the undefeated #1 North Carolina team, over #2 Kansas (and its seven foot center Wilt Chamberlain), was quickly followed by two more ultimate contests in 1961 and 1962. Thankfully, the tournament draw format (that was employed at that time) didn't force those two contests to be scheduled within one of the four regions (like in 1976) because the top two team's (Ohio State and Cincinnati) campus locations are separated by less than 150 miles. Geographically, both teams would probably have been invited to play in the Mideast region, except that the Missouri Valley conference champion (Cincinnati) had always been invited into the Midwest region, thereby allowing these Ohio rivals to postpone their confrontation until those highly anticipated, championship games. (As previously mentioned, the Big 10 conference champion, Ohio State in this case, was placed into the Mideast region.)

The 1965 finale showcased #1 Michigan against #2 UCLA, however, it wasn't until 40 years later (2005) that an audience was treated to another game between the consensus two best teams: Illinois and North Carolina. Why did it take so long for this to happen again, i.e., what was precipitating this situation to occur less frequently, or, were those four previous contests (in nine years, from 1957-65) an anomaly? (As highly desired as #1 playing #2 in the NCAA championship game is, even more interest would occur with a contest between two undefeated teams. Only three times have two teams with zero losses been invited to the tournament, and one of those teams lost their second tournament game in both 1968 and 1973. However, the 1976 Final Four had two such teams, and Michigan 'upset' Rutgers, disappointing those who were anticipating the ultimate final game that year: Rutgers against Indiana - with zero losses between the two of them.)

## Bracket Pairings: Before and After Seeding the Teams

Because of the published, pre-season pairings within the tournament draw, it was possible that the eventual top two teams in the polls would be aligned to meet in a Final Four, semi-final game, rather than the championship game. Thankfully, that particular situation only happened in seven of the 30 years beginning with 1949. There were five exceptions during those 30 years: 1953, when the #2 team accepted an invitation to the NIT; 1954, when #1 Kentucky declined their NCAA invitation because their graduate student athletes were ruled ineligible for postseason play; 1969 and 1973 when the #2 teams were on probation each of those years; and finally, in 1976, when the #1 and #2 teams met in the Mideast regional final, because of their campus locations, something that most assuredly would be avoided when employing current bracket assignment practices. After 1978, this possible misalignment has occurred 12 times in the last 33 years, with five semi-final games actually pitting the top two teams in the polls against one another. The year 2004 marked the first time when the selection committee ranked the four #1 seeds in an effort to postpone their top #1 seed playing their second best #1 seed, until the last tournament game. (Who the committee deemed the best two teams may not have always agreed with who the polls identified as the best two teams.)

With the possibility of inviting conference runner-ups to the tournament beginning in 1975, all top 20 teams (in the polls) have been present in the NCAA tournament ever since (unless prohibited from competing for probationary reasons). Before 1975, it was possible that one or more top teams were absent, like in 1974, when #4 Maryland could not be invited since the #1 NC State team (who won the NCAA tournament that year) earned the ACC's automatic bid by defeating Maryland in the conference's postseason tournament championship game. However, the number of top four teams in the Final Four (FF), according to the polls, was not essentially the same in the three tournament period groupings that are listed in Table 1 below.

Table 1 – History of the top 4 teams reaching the Final Four.

| Time Period | 1951-74 | 1975-84 | 1985-2011 |
|---|---|---|---|
| # years | 24 | 10 | 27 |
| # invitations/year | 16-25 | 32-53 | 64-68 |
| 0 top 4 teams in FF | 1 | 0 | 3 |
| 1 top 4 team in FF | 3 | 4 | 10 |
| 2 top 4 teams in FF | 12 | 6 | 9 |
| 3 top 4 teams in FF | 8 | 0 | 4 |
| 4 top 4 teams in FF | 0 | 0 | 1 |

Perhaps the parity of current team's abilities could explain the shift to having only typically just one or two of the top four teams reaching the FF. However, it is probably more likely that inviting a larger number of teams to compete in the NCAA tournament (especially since 1985), thereby forcing teams to play more games before reaching the FF, increased the likelihood that the top four teams would be beaten before reaching the FF. Therefore, it is not surprising that two or three of the top four teams competed in the FF more often before 1975, even though not all top four teams were always invited during the early years. Only two top four teams were invited in 1953 (one team was upset), in 1954 (when neither made the FF), and in 1959 (with

both teams reaching the FF). Nine years had only three of the top four teams invited: twice all three reached the FF, six times just two did, with one team making it the other year. In the 12 other early years, only once did just one top four team earn their way into the FF, with the other 11 years evenly split between two and three teams reaching the FF.

## Which Teams Deserved an Invitation?

To answer the question of how evenly balanced each region has been in the three, different bracket periods, a reasonable methodology had to be established so that the performance of each team in a given year could be quantitatively derived. Those three bracket periods are: 1950-1974, where only one team per conference could be invited; 1975-1984, which initially expanded the field to at most two invitations per conference in 1975, and then lifted that restriction in 1980; and finally 1985 to the present, when the field expanded to 64+ teams.

Previous experience has shown that there is no perfect system and many researchers have argued about the merits of rating systems versus ranking systems. The former typically incorporates the final score, or perhaps introduces some margin of victory (MOV) cap, or possibly applies a compression-like function to the full MOV, in an attempt to determine each team's rating as a measure of how many points better they are than the other teams. However, ranking systems cap the MOV to be at most one point, altering the focus simply to just wins and losses (and tie games - if allowed), thereby attempting to determine which teams have achieved the most outstanding (or meritorious) seasons, given their list of conquests and defeats. Therefore, the approach taken here will be more of a consensus strategy, and it bears a strong resemblance to how the Bowl Championship Series (BCS) attempts to determine who the best two NCAA football teams are, in a given season, so that they can play against each other in the designated BCS National Championship game.

The two primary components used in this approach will each contribute equally (50%) to the Tournament Selection Ratio (TSR). The first component will be an objective quantity that will be derived from eight computer-based systems. Four of these systems have gained wide recognition, regarding their merits, and, they were included since the results that they produce can be obtained - or generated - fairly easily. The other four systems have been created by this author, and represent somewhat different strategies for evaluating a collection of teams (relative to each other).

Four objective ranking systems were selected (all of which ignore MOV) and four more rating systems, three of which utilize the full MOV, complete the set of eight systems. Using a modified Borda counting strategy, each of the systems will generate their respective ratings/rankings, and the top teams (according to each system) will be assigned a 'vote tally' of 70, the #2 teams receive 69, and so on down to the $70^{th}$ top teams earning a vote tally of 1. (More about why only the top 70 teams receive any votes will be explained shortly.) The highest and lowest vote tallies are both ignored, and the remaining six vote tallies are added together, then averaged, and this value is finally normalized by dividing it by 70. This trimmed mean of the Borda count is then halved, making up 50% of the TSR. (It seemed wise to include a variety of systems in this objective component, in hopes that each team's true capabilities could be accurately measured - just from their game scores that year. It also seemed that each system was

just as worthy *a priori* as any other one included here, so each system's vote was treated equally when incorporating them all into this Borda count variation.)

The other 50% included in the TSR comes from a less objective source: the writers' and coaches' polls. Now even though some objections could be raised about the accuracy/validity of these polls, a case can be made where short duration injuries may have impacted a team, and an expert that has seen them play may be able to understand that they are better than what their record, or 'body of work', that year has been numerically reduced to (via either style of quantitative evaluation system). The number of votes a team receives in each poll will also be normalized (dividing by the maximum number of votes possible), and then those two values will be divided by four, so that each poll contributes equally to the TSR: 25% apiece.

The power rating system (PW - Carroll, et al, 1998), is included as one of the four rating systems as well as one of the four ranking systems (P1) by limiting the MOV to be at most one point in the latter case. The Rewards system (RW - Trono, 2007) is another ranking system, while a third such system (RP) is simply the **original** Rating Percentage Index (RPI) formula (as used by the NCAA selection committee - from the early 1980's to 2005) where: a team's won-loss record accounts for 25% of their RPI; their opponents' winning percentages contributes another 50%; and finally, a team's opponents' opponents' records comprise the final 25%. (The updated RPI formula, as employed by more recent NCAA tournament selection committees, weights road wins more than wins earned at a neutral site, which are weighted more than wins played in front of one's home fan base.) The fourth ranking system (MP), and two other rating systems (ED and SD) are described in more detail in Appendix A (along with a brief synopsis for RW).

The fourth (and final) rating system is the one invented by Jeff Sagarin, whose ratings have appeared in the daily *USA Today* newspaper since 1985. (Unfortunately, this implies that when calculating the TSR before 1985, the computer-based component will be derived from only the five values that remain - once the lowest and highest vote tallies have been dropped.) Anyone who has visited the web page that lists current NCAA Sagarin ratings might have noticed that two other columns list the ELO ratings, i.e., the ranking version of this system, and the Predictor ratings, which use scores more in line with the actual MOV. The listed Sagarin rating itself is some (non-disclosed) combination of these other two systems.

Even though the TRS does bear a strong resemblance to the BCS formula, hopefully, how the TSR is used here will not generate as much controversy as the formulation and components included in, and/or results generated by, said BCS formula. The primary purpose for the construction of the TSR was to have a plausible, quantifying methodology to evaluate the relative strengths of teams competing in the NCAA basketball tournament's four regions. Therefore, one way to validate how reasonable it is, as a strategy to rank teams, would be to compare the TSR with some strategies that have attempted to predict which teams would be invited to the NCAA tournament as non-automatic qualifiers. Once that is quickly examined, the bracket equality results, after applying the TSR, can be disseminated and examined. (The TSR can also be used to determine if the assigned tournament seeds have also been reasonable as well as evaluating if certain regions, which may vary from one year to the next, have had more, or fewer, really strong teams invited to compete in them, in comparison to the other regions.)

## Determining At-Large Teams

Though it was designed for a somewhat different purpose, the TSR has performed quite well with respect to matching who the NCAA selection committee has awarded the remaining, non-automatic invitations for the NCAA tournament. With regards to what other authors have proposed, TSR is comparable to the models that several researchers have trained using previous data, before evaluating their models on subsequent years of the NCAA tournament selection process. A brief summary of such a comparison is presented here, to help validate the TSR formula as a reasonable approach for ranking teams, according to their performances, in a given year. (A more detailed analysis describing how the TSR has ranked teams that were purportedly either invited serendipitously, or were excluded from the NCAA tournament, can be found in Appendix B.)

As can be seen in Table 2, applying the TSR to predict who the invited at-large teams will be compares favorably to the results produced by the dance card (DC) model devised by Coleman and Lynch (2001). In fact, the TSR's performance in this regard is even more impressive given it was created without using any training on any data sets. The DC model (based upon probit analysis), used the at-large bids in the 1994-99 tournaments to determine the weights to be associated with the six team attributes included in that model (RPI rank, number of wins over top 25 teams, conference (win-loss) differential, etc.), and then the model's performance was evaluated using the 2000 tournament. TSR correctly chose 33 of the 35 at-large teams in 2000, and the DC model chose 32.

Table 2 - System performance predicting NCAA tournament at-large bids.

| Span | TSR | PW | P1 | RP | RW | EX | MD | SD | SG | DC | All |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 01-11 | 340 | 307 | 346 | 338 | 324 | 302 | 332 | 288 | 337 | 351 | 377 |
| 94-00 | 218 | 185 | 215 | 212 | 207 | 187 | 216 | 179 | 205 | 224 | 240 |
| 85-93 | 277 | 251 | 275 | 270 | 260 | 238 | 269 | 230 | 265 | ----- | 306 |
| Total | 835 | 743 | 836 | 820 | 791 | 727 | 817 | 697 | 807 | 575 | 923 |
| Pct. | 91 | 81 | 91 | 89 | 86 | 79 | 89 | 76 | 88 | 93 | 100 |

The DC model has been modified, as described in Coleman, DuMond and Lynch (2010), to include a slightly different set of team attributes as well as information about each team's conference membership, and if there was any representation on the NCAA selection committee related to that institution, or conference affiliation. This updated DC used data from 1999-2008 to determine its weights, and then its effectiveness was evaluated using the 2009 tournament in that paper. The DC column in Table 2 was determined using the associated web site, http://www.unf.edu/~jcoleman/dance.htm, which provided the results for the unbiased model outlined in the 2001 paper. (The updated DC is slightly more accurate than the original.)

Table 2 also illustrates how effective each of the eight computer systems (which make up the TSR) is individually, with regards to predicting which teams will be invited. The systems which ignore margin of victory (P1, RP, RW, and MD) are significantly more accurate than PW, EX and SD. (You may recall that the Sagarin Ratings (SG) are an undisclosed combination of the ratings generated by his margin of victory system, aka Predictor, with the results produced by the

one that ignores margin of victory, i.e., his ELO system.) Using just P1, RP, RW and MD - to compute a modified TSR - did not provide any significantly improved, overall TSR performance in this 27 year span, though there were several disagreements during this time period, that were essentially negated in other years. (Ignoring the two polls, the computer portion of the TSR was less accurate, missing seven more at-large selections than the full TSR.) Reilly (2003) used components similar to those in the original DC model, correctly selected 32 of the 34 teams in 2003 - as did the DC; TSR had 33 that year, only missing North Carolina State's selection.

Using a different implementation strategy, Jing and Cox (2008) employed a neural network approach to solve this selection problem, while also investigating how accurately the seeds have been assigned. Their three models were only evaluated with regards to the 2005 tournament, and correctly selected 31, 32 and 33 teams respectively (out of 34 at-large selections in 2005), while TSR correctly chose 31 teams, and the original DC model matched 33 at-large selections.

If the original intention behind the creation of the TSR had been to create a system to match as many of the at-large teams invited to compete in the NCAA tournament as possible, then it would have been prudent to calculate appropriate weights, in a manner similar to Coleman and Lynch, for each of the 10 individual TSR contributing polls/systems, to maximize the number of correct matches in a modified TSR calculation. As stated previously, that was not the rationale behind the creation of the TSR. However, it is interesting to note that the weights in the original DC model **were** computed to maximize that total (from 1994-99), and so it is somewhat impressive that the TSR correctly would have predicted 185 such invitations, to the DC model's total of 192, in that six year, training period. Upon further examination of the quantities in Table 2, one can easily observe that the P1 system has been the best individual predictor of which at-large teams will get invited, and by itself is almost as accurate as the trained DC model!

## Bracket Evaluation

The previous section attempted to justify the merits of the TSR by comparing its effectiveness, regarding the ranking of teams, with strategies that predict which teams will receive an at-large NCAA tournament bid. That ranking feature of the TSR will now be used to quantify how evenly balanced each of the four NCAA tournament regions have been in the past.

To compare the relative strengths of the teams that are placed into each region, a strength function was required that would be low for teams not appearing near the top of the TSR list, and, it would generate increasingly larger values for teams that have demonstrated a strong chance of emerging as the NCAA champion. There have been at least 64 teams competing in the tournament since 1985, and the aforementioned value of 70 was selected as the cut off because: it is larger than 64, and, $70^2 + 100 = 5000$, which is a nice round number (for the #1 team) - and so is the value of 70. Using the formula $(71 - \text{TSR rank})^2 + 100$, for teams ranked in the TSR top 70, the $70^{th}$ team is assigned a strength of 101, the $69^{th}$ team 104, $68^{th}$ team 109, etc., up to 4861 for the #2 team, and the maximum strength value of 5000 for the #1 team. (Adding any positive value, V, to each strength value dramatically reduces the relative difference between the #1 and #70 teams, since that quantity would be computed as $(V + 4900) / (V + 1)$. If V is 100, this yields 5000 for the #1 team, and that seemed to be a good choice that was also relatively close to 4900.)

As mentioned previously, the 2004 selection committee ranked the four #1 seeds, and it also began using the 'seeding S-curve' as well. The top #1 seed was paired with the lowest #2 seed (the #8 team overall, in the minds of the selection committee) to compete in the same region, and so on, with the top #2 seed (the #5 team overall) in the same region as the #4 team overall (which is the lowest #1 seed). This continued into the #3 and #4 seeds as well, thereby placing the top #3 seed into the region with the lowest #2 (and #4) seed. (There are some restrictions when seeding teams, but those are outlined in the NCAA Principles and Procedures link in the references.)

If the strength function is applied to the top 16 teams in the TSR ranking, and they are seeded according to this S-curve, then the total in one region (with the #1, #8, #9 and #16 teams) would be 5000 + 4069 + 3944 + 3125 = 16,138. The totals in the three other regions would be 16,114, 16,098 and 16,090 respectively. However, the selection committee's evaluation of the teams will not follow the TSR's ordering, and so, the regions' totals will be more disparate. If the best 64 teams were in the tournament, and the S-curve strategy was continued from the #1 seeds down to the #16 seeds, then the average region total would be 30,626. When examining the average total strength per region, from 1985-2011, the value was slightly lower: 29,199.1. The primary reason for this discrepancy is that almost all teams seeded #15 or #16, and some teams seeded #13 or #14 in a region, do not appear in the TSR top 70, and therefore, those teams would contribute zero to these totals, thereby producing lower averages.

What would constitute an acceptable deviation from the average regional strength value? To determine this, a reasonable, worse case analysis will be performed. If the best seeds (#1-#16) were all in one region, and the worst seeds were in another, then those two regions' totals would differ by 3,696: 411 (5000-4589) + 387 (4456-4069) + 363 + 339 … + 75 + 51 (200-149). (Though this would not necessarily be the most appropriate alignment of the seeds into the four regions, it would have at least met the bare minimum criteria that one team with each specific, assigned seed would have been placed into each region.) In the 27 years from 1985-2011, the difference between the largest and smallest region total strength values was less than two thousand three times, and less than four thousand 11 times, which implies for those years, using the TSR strength formula, that the committee did a fairly good job of placing teams into the four regions in a balanced manner. (This does not validate that the best teams were selected, nor were the seeds assigned to these teams fairly; it only indicates that the overall quality of all the teams in each region were roughly equivalent, as defined by this quantitative approach.)

It could be argued that the total strength values of two regions differing by close to 3700 implies that those two regions are unbalanced, as that difference is larger than 10% (of the average, regional strength total). Eight times (in the last 27 years) the largest difference was between four and six thousand times, and it was between six and eight thousand four more times, and only once was it larger than that. The years with the two largest differences were examined more closely, and here is what was uncovered. The second largest difference occurred in the 2007 tournament; the Midwest region's total (31,198) was roughly seven thousand more than the East region's total (24,103). There were many instances where the TSR would have seeded teams differently within those regions, and in this particular case, swapping the #6 seeds in these two regions reduced the difference from 7095 to 1927, creating four fairly evenly balanced regions.

The tournament with the largest difference discovered was played in 1988, and the West region's total was over 10,000 more than the Midwest region (35,381 versus 24,838). The difference was only about two thousand when summing the strengths of the top four seeds in each region, and TSR's top four teams were the #1 seeds in the four regions. According to the TSR ranking, 16 of the top 20 teams were the #1-#4 seeds, but the Midwest had the TSR's #3, #15, #13 and #19 teams, which helps to explain the two thousand point difference with the West, which included TSR's #2, #7, #10 and #11 teams. The teams in those regions were even more different in the next four seeds: the Midwest region included #25, #33, #31 and #64, whereas the West had #17, #24, #14 and #37. The West also had TSR's #16 team (Loyola Marymount) as its #10 seed. (That TSR ranking is somewhat justified since Loyola did defeat the #7 seed in that region that year.) Swapping Loyola to the Midwest, and doing likewise with the #4, #7 and #8 seeds brings the difference down to almost zero, so moving three of these four teams would make those two regions - along with the other two - more in balance. (If NCAA tournament seeds were allocated in order, according to the TSR rankings, those seeds would differ by 1.056, on average, with the NCAA tournament selections committee's choices from 1985-2011. Over that 27 year span, each tournament has roughly 24 teams whose assigned seeds are validated by the TSR, another 24 whose actual seed is plus or minus one from what that team's seed would be, using the order specified in the TSR ranking, another 9 teams that are off by two, and 7 more teams where their tournament seed is off by more than two, as would be derived from the TSR ranking.)

So, if shifting only a few teams creates a reasonably balanced set of regions, according to this analysis, then perhaps the selection committee has done that part of its jobs fairly well - at least since the field was expanded to include 64 teams. As it turns out, sports announcer Billy Packer did make a few pertinent comments near the end of the Selection Sunday TV show back in 1988, concerning the main focus of this report. He thought that Purdue (#3 in both polls heading into the NCAAs) had the best chance of reaching the Final Four that year since the region where it was assigned as the #1 seed (Midwest!) was in his opinion the weakest. As it turns out, the #6 seed in that region (Kansas) emerged as that region's representative in the Final Four, and they were eventually crowned as National Champions that year.

The 1988 tournament will come up again shortly, but it may not be so easy to quantify how balanced the regions were before 1985, when there were fewer invited teams, especially before 1975 - when only one team per conference could be invited, and, when the number of teams invited to each specific region varied from year to year. As previously mentioned, selection committees before 1975 were locked into placing teams into regions geographically because their choices were severely constrained by the published tournament draw template that they were obligated to adhere to. Al McGuire (NBC sports announcer - and head coach at Marquette before taking that job) once commented during a televised game, about the tournament regions after he joined NBC.

McGuire said that he was always hoping that his team would be invited to compete in the Midwest region instead of the Mideast region since the Big 10 and SEC conference champions were always present in the Mideast, and those conferences were perceived (at least by coach McGuire) to be superior to the (then) Big 8 and Missouri Valley representatives (in the 1970s) that were invited to compete in the Midwest region. (Marquette was invited to play in the Midwest region in 1977 for only the second time - 1961 was the first - instead of competing in

the Mideast region. Marquette went on to become the National Champion that year, in McGuire's last year of coaching.) Other sports journalists have made similar comments, criticizing the lack of strong competition for UCLA in the West region, citing that as one possible reason why UCLA was able to be crowned as the National Champion 10 times during the years from 1964 to 1975. (Investigating comments like these are the major reason why this study was attempted!)

To continue investigating the question of bracket equality, a related but slightly different methodology must be introduced to help evaluate how strong each region was, with respect to the teams contained therein. This methodology can be applied to the tournaments with 64+ teams as well, but first a model, for predicting which teams in a region are more likely to advance to the Final Four, must be described.

## Probability of Reaching the Final Four

Many researchers (Brown and Sokol, 2010, Coleman and Lynch, 2009, and West, 2008) have created a variety of computational models for predicting NCAA tournament results: from determining probabilities for individual game outcomes, to calculating the probability that each team will advance to the Final Four, or even become the tournament champion. Some of these models use each team's ratings, or other objective, quantitative measurements, to determine these likelihoods. However, the straightforward model (described below) relies solely on the assigned, regional seed # of each team.

Breiter and Carlin (1997) used the following formula to approximate the probability that a team with seed n would defeat a team with seed k: k / (k+n). They then calculated the expected probability that each seed would reach the Final Four. Berry (2000) modified this strategy to use a seed's strength, where the strength values were mathematically determined so that they would minimize the sum of all the squared error terms when compared against the observed results in the NCAA tournaments from 1985 to 2000. (The probability of seed n defeating seed k is: $Strength_n / (Strength_n+Strength_k)$.)

To evaluate a team's chance to reach the Final Four, in the context of this study, the TSR strength was modified slightly before being incorporated into Berry's formula. The strength of each team in the TSR's top 70 would be $(71 - \text{TSR rank})^2 + 100$, therefore, the #70 team is assigned a strength of 101. Since many teams in the tournament (typically from 11 to 19) are not in that group, the next 99 teams in the TSR ranking were assigned strengths equal to 171 – TSR rank, so team #71 has strength 100, #72 99, #73 98, and so on down to team #169 being assigned a strength of 2, where all other lower ranked teams are assigned a strength of 1.

Using the formula, where the probability that team n defeats team k is $TSRstrength_n / (TSRstrength_n+TSRstrength_k)$, the probability of each team reaching the Final Four can be calculated (as done in Berry, 2000). From 1985-2011, most #1 seeds have had a 24-30% chance to reach the Final Four using this model, and the #2 seeds were in the range 19-25%. Teams with a high TSR ranking, who are also in a weak region, will have a higher expected likelihood of reaching the Final Four than average. In fact, according to this model, Purdue had a 35% chance of reaching the Final Four in 1988, given the weak region it was in.

Table 3 compares the likelihood of each seed making its way through the first four rounds (ignoring all 'play-in' games when more than 64 teams are invited) for the actual TRS Strength values from 1985-2011 versus the Strength values that would be assigned if the top 64 teams were invited, and the first four were #1 seeds, the next four #2 seeds, and so on down to the last four invited being #16 seeds. (In 15 out of these 27 years, the top four TSR ranked teams have been the four #1 seeds. It is also true that on four occasions, all four #16 seeds have been ranked as the #170 team or lower, primarily because of the weaker conference representatives that are automatically invited, and who are typically designated as the #16 seeds.)

Table 3 - Probability of reaching the Final Four.

| | Actual Avg. | Prob. | Top 64 Avg. | Prob. |
|---|---|---|---|---|
| 1 | 4768 | 28.10 | 4793 | 26.69 |
| 2 | 4185 | 22.94 | 4261 | 21.71 |
| 3 | 3678 | 17.26 | 3761 | 16.75 |
| 4 | 3212 | 12.60 | 3293 | 12.26 |
| 5 | 2679 | 7.27 | 2857 | 8.59 |
| 6 | 2310 | 5.08 | 2453 | 5.80 |
| 7 | 1831 | 2.60 | 2081 | 3.66 |
| 8 | 1655 | 1.85 | 1741 | 2.14 |
| 9 | 1367 | 1.03 | 1433 | 1.21 |
| 10 | 1189 | 0.71 | 1157 | 0.66 |
| 11 | 892 | 0.31 | 913 | 0.33 |
| 12 | 829 | 0.26 | 701 | 0.15 |
| 13 | 348 | 0.02 | 521 | 0.06 |
| 14 | 209 | 0.00 | 373 | 0.01 |
| 15 | 59 | 0.00 | 257 | 0.01 |
| 16 | 21 | 0.00 | 173 | 0.00 |

The actual averages of the TSR strengths assigned per seed from 1985-2011 are very close to what the averages for the top 64, TSR strength values would be. One would expect the top 64 averages to be higher, because the actual average includes the invited teams that are below the first 64 teams appearing in the list ordered by the TSR rating, but the #10 and #12 averages were slightly higher, possibly indicating some teams were seeded incorrectly. The disparity between the averages for the #5 and #7 seeds was close to a couple hundred, and all the other seeds' averages were fairly close, except the #13 to #16 seeds, where the averages were significantly lower, relatively speaking, because of all those automatic qualifiers from weaker conferences.

To include more than just the top seeds in the following analysis, the TSR Strength model's ability to predict the likelihood of the seeds reaching the Elite Eight is summarized in Table 4. The actual number of teams, for each seed that reached the Elite Eight, is included along with the predicted number to do so, using the top 64 average strength values. This is contrasted against the slightly more accurate predictor of the actual, average TSR strength values used during that 27 year span.

Table 4 - Expected number of teams to reach the Elite Eight: 1985-2011

| | Actual | Top 64 | TSR Str. | Berry | B (Str) |
|---|---|---|---|---|---|
| 1 | 78 | 49 | 52 | 76 | 100 |
| 2 | 53 | 43 | 45 | 51 | 43 |
| 3 | 25 | 35 | 36 | 24 | 25 |
| 4 | 14 | 27 | 28 | 14 | 25 |
| 5 | 7 | 20 | 18 | 9 | 21 |
| 6 | 12 | 15 | 13 | 15 | 21 |
| 7 | 6 | 10 | 8 | 8 | 17 |
| 8 | 7 | 7 | 6 | 4 | 17 |
| 9 | 1 | 4 | 4 | 4 | 17 |
| 10 | 7 | 3 | 3 | 5 | 14 |
| 11 | 5 | 2 | 2 | 3 | 10 |
| 12 | 1 | 1 | 1 | 2 | 10 |
| 13 | 0 | 0 | 0 | 0 | 6 |
| 14 | 0 | 0 | 0 | 1 | 5 |
| 15 | 0 | 0 | 0 | 0 | 2 |
| 16 | 0 | 0 | 0 | 0 | 1 |

Berry's model, where the seed's strengths were calculated to approximate the observed behavior as closely as possible, matches the number of teams to reach the Elite Eight more closely than the other two models. However, the strength values employed in Berry's model (which can be found in the rightmost column in Table 4) are more dramatically separated in some instances, e.g., the strength of a #1 seed is more than twice the strength of a #2 seed, which is almost double the strength of a #3 - while other values are identical, e.g., the strength of the #3 and #4 seeds; likewise with seeds #5 and #6, and seeds #7, #8, and #9 as well. The TSR strength formula incorporated more gradual decreases with each increasing seed value, and *a priori*, this seemed to be a more reasonable manner in which to construct that formula.

Using the TSR strength formula, the average probability that a #1 seed would reach the Final Four was roughly 27% over the last 27 years, and roughly 22% for #2 seeds, with a standard deviation (SD) equal to 0.03 in both cases. There were four #1 teams with probabilities greater than the mean plus two SDs (and one below the mean minus 2 * SD) and three #2 seeds above (and another #2 below) that same boundary. Table 5 summarizes information about those instances.

Table 5 - Outliers with regards to the likelihood that a team will reach the Final Four

| | | Year | Team | Str.V. | Region | Prob. | Top Region Opp. |
|---|---|---|---|---|---|---|---|
| + | 1 | 1986 | Duke | 5000 | 30320 | 34.39 | 9, 14, 16, 22 |
| + | 1 | 1988 | Purdue | 4724 | 24837 | 35.09 | 9, 13, 19, 25 |
| + | 1 | 2001 | Duke | 5000 | 25304 | 36.26 | 9, 7, 14, 28 |
| + | 1 | 2002 | Mary. | 4456 | 26184 | 34.17 | 10, 23, 16, 11 |
| + | 2 | 1991 | Indiana | 4861 | 29593 | 29.00 | 1, 10, 27, 48, 35 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| + | 2 | 1996 | Cinc. | 4589 | 27179 | 30.73 | 3, 15, 29, 45, 48 |
| + | 2 | 2003 | Pitt. | 4724 | 30379 | 29.26 | 1, 11, 24, 48, 39 |
| - | 1 | 2005 | Wash. | 4069 | 30053 | 18.09 | 3, 11, 4, 25 |
| - | 2 | 2006 | Tenn. | 2909 | 27611 | 14.71 | 3, 9, 27, 34, 58 |

The rightmost column contains the TRS rank for the possible opponents in the last two/three regional contests (for each team listed): the #2 - #5 seeds for a #1 seed, and, the TSR rank of the #1, #3, #6, #7 and #10 seeds with respect to the #2 seeds. The 1986 Duke team's high Final Four probability (FFP) is due to its #1 TSR rank as well as being paired with the weakest #2 and #3 seeds that year (as determined by the TSR). Likewise in 1988, Purdue had the weakest #2 and #5 seeds in its region along with the second lowest, overall total ever in a region (as measured from 1985-2011). The 2001 Duke team (once again) was awarded the highest TSR ranking and was placed in a very weak region overall (except for a strong #3 seed) which led to the highest FFP observed so far. (A similar case can be made for the 2002 Maryland team, and their region: both the 2001 Duke team and the 2002 Maryland squad were crowned national champions.)

The 1991 Indiana squad was a #2 seed, but TSR had them as its overall #2 team, so it is not unexpected that their FFP was a high outlier since the Hoosiers would have been awarded the second highest TSR strength value. The #2 seeded, 1996 Cincinnati team really only had one strong, possible opponent to play in its last three, regional NCAA tournament games, which helps to explain how it was assigned such a high FFP; the 2003 Pittsburgh team's case as a #2 seed is similar. (Cincinnati was TSR's #4 team overall, and Pittsburgh was #3.)

As far as the underachievers in Table 5 are concerned, Appendix B explains in a little more detail about the 2005 Washington team; suffice it to say that the #2 and #4 seeds in their region had higher TSR rankings than the Huskies. The 2006 Tennessee team was awarded a #2 seed but they were only regarded as the #18 team in the TSR ranking, which significantly impacted their FFP (even though they only had two strong opponents in their region that they could have played in their last three regional, tournament games).

With these probabilistic models now in place, the years before 1985 can be quantitatively analyzed/scrutinized, especially before 1979, when seeding began, and the preseason tournament draw template was eliminated.

## The Tournament Draw's 'Unfairness'

Several questions have been previously posed in this essay, with regards to how equitable the tournament draw mechanism was, concerning the balance in each region, and the results of that investigation will be provided shortly. One other item to consider is that the tournament draw format also occasionally made it more difficult for some teams in a region to not only be crowned the tournament champion but also just to reach the Final Four.

Unless many upsets occur in the early tournament rounds, the team that earns the title of champion will most assuredly need to defeat several other top teams during the tournament, so in one sense, should it matter if the eventual champion defeats the #1 (and/or #2) team in some early round, or in the championship game? It seems like the answer should be no, but how many

top teams should the champion have to defeat and still have the tournament alignment be deemed fair?

For instance, let us examine the format for a 32 team tournament that was seeded according to the aforementioned procedures used to create a 32 player tennis tournament. The #1 seed should play the #32 seed, followed by a match against the #16 (or #17) player, and then the #8 or #9 seeded player, unless upsets have eliminated both of the latter two players.

The 1976 Indiana Hoosiers are the last team to complete the **entire** season undefeated. (Only six prior teams have accomplished this - culminating with winning the NCAA tournament.) Voted #1 in both polls, the Hoosiers certainly would have been recognized as the overall #1 seed, yet because of the tournament draw template in place back then, their first opponent (St. John's) was #18 in the final, regular season coaches' poll. After defeating them, Indiana's next opponent (Alabama) was ranked #6 in the AP pool, and #7 in the UPI poll - which is quite a distance from #16 or #17. Their third opponent (Marquette) was the #2 team in both polls - and this game was only the NCAA Mideast regional final!

It does appear that by placing teams in the regions by geography/conference affiliation, the tournament draw was unfair to Indiana in the sense that their path to the Final Four was at first glance more difficult than teams in other regions that year. To quantify this concept of fairness, if points were awarded for wins over ranked opponents as follows, then one could calculate a measure of the quality of opponents faced in the tournament. Since typically over 40 teams receive votes in both polls, $(51\text{-TSR rank})^2$ will be the number used as fairness points (FP) for defeating a ranked team. (This value would be the average of the two FP values if the rank in both polls were unequal, and zero for teams that did not receive any votes in either poll.) Table 6 lists teams that had to defeat many highly ranked opponents during the tournament on the way to earning their championship.

Table 6 - Evaluation of fairness reaching the NCAA championship game.

| Year | Champion | Fairness Points | # G | Total Games | Opponents (in order → 1st to last) |
|---|---|---|---|---|---|
| 1985 | Villanova | 7063.5 | 3 | 5 | -,2,-,7,5/4,1 |
| 1950 | CCNY | 7017 | 3 | 3 | 2,5,1 |
| 1997 | Arizona | 6825 | 3 | 6 | -,16/18,1,-,4/5,5/4 |
| 1983 | NC State | 6734 | 3 | 6 | -,6,4,18/15,1 |
| 1985 | Villanova | 8999.5 | 4 | 5 | (See above) |
| 1963 | Loyola (IL) | 8864 | 4 | 5 | -,6/7,8/5,2,1 |
| 1976 | Indiana | 8193 | 4 | 5 | -/18,6/7,2,5,9 |
| 1996 | Kentucky | 8068 | 5 | 6 | -,22/21,12/10,9,1,15/14 |
| 1966 | UT El Paso | 8041.5 | 5 | 5 | *-/13,7/9,4,-/12,1 |
| 1983 | NC State | 7926.5 | 5 | 6 | (See above) |
| 1980 | Louisville | 3377.5 | 5 | 5 | -/20,-/18,3/2,-,- |
| 1967 | UCLA | 2593 | 4 | 4 | -,-/16,7/6,- |
| 1954 | LaSalle | 1506.625 | 5 | 5 | -,10/14T,-,-,- |

No matter how you break it down, the 1985 Villanova team had the toughest road to its championship. The Wildcats' second round game was against the #2 team in the country, and their Final Four opponents were #5(AP)/#4(UPI) and #1 respectively. For these three games, their FP total would be 7063.5, which increases to 8999.5 when you include their fourth highest rated NCAA tournament opponent (the #7 team). The only team to win both the NIT and NCAA tournaments (City College of New York - CCNY) had the second highest total, over three NCAA tournament games (and the highest FP average including all tournament games, given only eight teams were invited in 1950). The 1997 Arizona team is the only one (so far) to defeat three #1 seeds, and they earned the third largest, three-game FP total. The aforementioned, 'Cinderella team' from North Carolina State had the fourth highest, three-game FP sum, and they were also the first NCAA champion who had to win six tournament games before being crowned.

The top three, four-game FP totals also appear in Table 6 as well as those teams that follow them when a fifth (or sixth) game is included, since those first three, four-game FP totals surpass the next three highest, five-game FP totals. (Indiana's five-game total would be slightly larger than its four-game total. Villanova and Loyola's totals would be unchanged - since their remaining, unranked opponents contribute zero FP - but both teams would still have the two highest totals. NC State would slip to sixth place, when considering a five or six-game FP total. Please note that dashes in Table 6's rightmost column indicate that the opponent in that game was not ranked.) The last three entries in Table 6 could represent the three teams having the easiest road to their Championship. Only 8 of 24 invited teams were ranked in 1954, including LaSalle, but #1 Kentucky refused their NCAA invitation (for eligibility reasons as mentioned earlier in this article). There is a small bias against #1 teams with this strategy since they can't gain 2500 FP by defeating the top team. However, a #1 who defeated other top 10 teams would still have a relatively high total, e.g., 1976 Indiana.

From 1953-74, the number of teams invited to the NCAA tournament varied between 22 and 25, and roughly 7 to 10 spots in the tournament draw were reserved for conference champions who were also awarded a bye into the next round. In those 22 years, only five champions were crowned after defeating five, NCAA tournament opponents even though more than half the invited field wasn't awarded a bye, which indirectly speaks perhaps to a different form of unfairness, with the previously employed, bracket filling strategy (known as the tournament draw).

## Comparing the Tournament Draw Versus When Teams are Seeded

Using the previously described tennis tournament model, where the better players are seeded so as to postpone playing other better players until later tournament rounds, these two NCAA tournament bracketing strategies can be evaluated. One methodology for performing this comparison would be to count the number of games where opponents should have been meeting further into the tournament. Table 7 summarizes these occurrences - where said contests were scheduled one round too early (in the tournament). The eight grouping categories (rows) were created for the following reasons: only 16 teams were invited in 1951-52; the number of teams invited varied between 22-25 from 1953-62; the AP poll only listed its top 10 teams from 1963-68; exactly 25 teams were invited from 1969-74; 32 teams were invited from 1975-78 (and conference runner-ups could be invited as well); the seeding of teams begins with 40, then 48, 48,

48, 52 and 53 teams respectively from 1979-84; 64 teams are invited from 1985-2000; and finally, 65+ teams are invited from 2001-2011.

As can be seen in Table 7, many games have been played which, according to the polls, should have been scheduled in the subsequent, tournament round. (No one is too worried about an unfair, first round pairing when #27 plays #31, but #21 versus #25, or #22 versus #17, is a more appropriate, second round matchup than an opening round game, at least with regards fairness - and how tennis tournament seeding works. Of course, the selection committee's ranking of teams may disagree with the rankings appearing in the polls, but as will be described shortly, the number of such discrepancies would probably not be to the same level observed as when the tournament draw was in place.)

Table 7 - Number of games where tournament pairings were one round too early.

| Span | 1st | 2nd | 3rd | 4th | 5th |
|---|---|---|---|---|---|
| 1951-52 | ----- | ----- | 1/16 | 1/8 | 0/4 |
| 1953-62 | ----- | 0/79 | 7/80 | 3/40 | 1/20 |
| 1963-68 | ----- | 2/45 | 12/48 | 3/24 | 2/12 |
| 1969-74 | ----- | 3/42 | 9/48 | 3/24 | 1/12 |
| 1975-78 | ----- | 11/64 | 2/32 | 1/16 | 0/8 |
| 1979-84 | 0/84 | 5/80 | 1/40 | 0/20 | 0/10 |
| 1985-2000 | 22/512 | 5/256 | 4/128 | 1/64 | 0/32 |
| 2001-2011 | 23/352 | 11/176 | 8/88 | 0/44 | 0/22 |

However, eight games were also played that could be considered occurring two rounds too early, at least if you agree with the teams' rankings (in the polls) - and with the tennis tournament pairing/seeding strategy: four of these games happened in the four years (1975-78) between when the cap of one invitation per conference was increased to be at most two, and when seeding began. (These eight games are not included in Table 7, and more specific grievances that occurred during these same four years will be described shortly.)

In the Midwest region, the #2 and #4 teams (in both polls - in 1957) were paired in their opening game (which corresponds to a game between Sweet Sixteen teams in present day tournaments), when according to the canonical tournament schedule, this should've been a Final Four semi-final game. The same situation occurred in 1958 as well, pitting the #2 against the #3/#4 team. In 1965 and 1968, the #3 and #4 teams played each other in a Sweet Sixteen game as well - both within the East region. Three of the other four games, which were played in the NCAA tournament between 1975 and 1978, are similar: #6/#4 versus #5/#6; #6/#7 versus #8/#6; and #3/#2 versus #4/#6 – all occurring in the first round (of 32 teams)! However, probably the most egregious such premature contest has already been mentioned: the 1976 Midwest regional final - pitting the #1 and #2 teams in both polls - which should only occur in the tournament's championship game, or at worst, a semi-final, Final Four contest (given current seeding practices).

Some notable trends accentuate the four year transition period (1975-78), when the evolution from the tournament draw to the present bracket equality strategy (of seeding) began. Because

many strong conference runner-ups could now be invited to the NCAA tournament, the observed occurrence of premature contests was higher even though only one game per year in the tournament draw format was predetermined, i.e., both teams in that pairing were specific conference champions. This means that the other 14 games had one at-large team pitted against a conference champion, since there was typically also one game where two at-large teams played against each other in an opening game. Therefore, the selection committee had great flexibility in aligning roughly half the field (teams receiving the at-large bids), against those specific conference champions that were placed into a specific bracket location (as listed in the published tournament draw). These latter placements were still geographically oriented, but the conference runner-ups, and other at-large teams, could still be placed into any region, though it appears that there was some sentiment to try and place teams close to their home region, as witnessed by which regions teams actually played in. Unfortunately for some teams, the selection committee appears not to have wished to set up the NCAA tournament like a similarly seeded tennis tournament - or at least not during the years from 1975-78.

Even though roughly two of the eight teams in each region would've appeared in a different region, according to its conference affiliation, as specified by previous tournament draw templates, two unranked teams played each other in the opening round (one team being a conference champion) when in the same region's bracket a #5/#6 at-large team opened its tournament up against a conference champion that was the #6/#4 team in the polls in 1975. This matchup was worthy of an Elite Eight contest, which is scheduled to occur two rounds further into the tournament; likewise when the #8/#7 team played the #10/#8 team in 1975 as well. Similar (stellar) opening round games also occurred in 1977: #10/#7 versus #7/#8 and #3/#2 versus #4/#6, the latter being a game worthy of Final Four status! Table 7 illustrates a preponderance of other games that were scheduled one round too early from 1975-78. In fact, there were more than twice as many such games, percentage-wise, in the opening round during that four year period than any other one listed: 17.2% in 1975-78, with the next highest value at 7.1%. Therefore, the (predetermined) tournament draw format facilitated more early, third round (Sweet Sixteen) games being played; there have been fewer such occurrences since the seeding process officially began in 1979, and these premature pairings have occurred only once since 1979 in the last two rounds of the tournament (that precede the championship game).

In 1978, the #1 Kentucky Wildcats opened up against Florida State (#15/#12) in the Mideast region, and if the #3 Marquette team hadn't been upset by Miami (Ohio), after their center was ejected for excessively swinging his elbows when grabbing a rebound in traffic - with Marquette comfortably ahead by about 10 points, with under 10 minutes to go in the second half - Kentucky would've played them in the Sweet Sixteen even though this game would have been worthy of the Final Four, which was still two rounds away! Why would the selection committee place the #3 team (an at-large team) right next to the #1 team in that region's bracket? At least, by putting them in the other half of that region's bracket, it would have postponed their possible matchup until the regional final (as occurred in 1976). And since the committee had so much flexibility in where at-large teams were placed, why not place this at-large team into the Midwest region where the #4/#7 and #10/#11 teams were placed? If Kentucky did have to defeat #3 Marquette, they would've surpassed Villanova as having the toughest road to the championship, amassing a 9795.5 FP sum - almost 800 points above Villanova's total - after including Kentucky's regional final win over the #4/#5 rated team. (This four game FP total for Kentucky would have been

slightly higher than Indiana's, which would have placed them in the third spot in Table 6. Of course, there are also other scenarios which would increase other teams' FP sums as well.) At first glance, Kentucky could possibly have had to defeat three top 15 teams just to win their region in 1978 - which in hindsight could be construed as purposeful, unfair bracket placement decisions by the committee that year, given all the different ways those at-large teams could have been rearranged by them. (Kentucky went on to defeat two more top 10 - #5/#6 and #7/#9 - teams in the Final Four that year as well.)

## Regional Balance in the Tournament Draw Era

Now that all the necessary, preliminary background work has been presented, one question that has been asked throughout this report can finally be addressed: how evenly balanced were the four regions when the tournament draw template was in use? The TSR strength value will be the primary measurement employed here, and because each region included between four and seven invited teams (from 1953 to 1974), before each region began inviting eight teams in 1975, the sum of the team's TSR strengths could be misleading. Therefore, to eliminate the impact of this varying number of teams per region, only the four teams with the largest TSR strength values in each region will be included, and that four-value sum will used to make some conclusions.

If a region has several strong teams, then the top team's TSR strength value should be around 30% of the overall total. (Using the specific values presented in the previous Bracket Evaluation section contained herein, i.e., the 16 largest possible TSR values, these ratios would 31.0, 30.17, 29.3 and 28.5 respectively for the four regions when the aforementioned seeding S-curve is applied.) Since the tournament began inviting 64 teams in 1985, the average ratio in each region has been: 30.39 (East), 29.79 (Southeast), 29.76 (Midwest), and 29.54 (West). However, these ratios are somewhat misleading because there is now not very much geographic focus when most teams are placed into one of the four regions by the selection committee, and that is quite the opposite of how the tournament draw previously locked specific conference champions into specific tournament slots, within specific regions.

Table 8 – Analysis applied to the Top 4 Teams in each region: 1953-1974

| | East | Mideast | Midwest | West |
|---|---|---|---|---|
| Minimum | 10350 | 10806 | 5836 | 7385 |
| Maximum | 17598 | 17485 | 15851 | 17094 |
| < 10,000 | 0 | 0 | 3 | 5 |
| 10,000-12,999 | 7 | 7 | 10 | 10 |
| 13,000-14,999 | 5 | 9 | 6 | 4 |
| 15,000-16,999 | 7 | 4 | 3 | 2 |
| > 17,000 | 3 | 3 | 0 | 1 |
| Average | 14414.2 | 14245.6 | 12428.9 | 11993.4 |
| Top 4 ratio | 32.51% | 33.50% | 37.20% | 39.34% |
| # > 40% | 1 | 2 | 6 | 11 |

Table 8 summarizes the results from studying the balance between regions, starting when the field was expanded from 16 teams (in 1953) up until the one team per conference restriction was lifted after the 1974 tournament. The average sums of the top four teams in the Midwest and West regions are clearly significantly lower than the East and Mideast regions, and there were many times when only one or two teams in the former regions appeared fairly close to the top of the TSR ranking, as indicated by the eight different times where one of those regions had a top four TSR strength sum less than 10,000. In contrast, the East and Mideast regions never seemed to suffer from that situation. In fact, they typically had three or four strong teams (as signified by a TSR strength value larger than 3125, the TSR strength value for its #16 ranked team) per region each year, as indicated by the large number of times where the aforementioned sum was greater than 15,000. The totals in the West region benefited from all those years where UCLA was dominant, however, even if the Bruins were the best team in each of those years, contributing 5000 points to those sums, it is still the case that in 6 of those 11 years, from 1964-1974, the West's top four sum was still less 13,000, and in four of those, the top four ratio - of the top team's TSR strength divided by the overall sum - was greater than 40%. The West region's ratio was above 40% seven of the nine years, from 1953-1961, and that is also where the five sums that were less than 10,000 occurred, indicating that the West region back then invited far fewer strong teams than the other three. (Two of the three years whose sum was less than 10,000, for the Midwest region in Table 8, also occurred during the 1950's.)

To even further highlight the discrepancy within (and across) those four geographically-based NCAA tournament regions, let us examine the ratio of the top two teams in each region, using those teams' TSR strength values. Table 9 lists the average ratios of the strength value associated with the top TSR ranked team in each region divided by the second highest team's strength value (in that region). This ratio is fairly low, and similar, for all four regions from 1985-2011. However, there is a noticeable increase in these values for the Midwest and the West regions when the tournament draw was in use. The largest ratio from 1985-2011 was 1.3574, when the second and twelfth highest TSR ranked teams were invited into the Midwest region in 2002. Contrast the fact that only two region's ratios were above 1.3 from 1985-2011 with the 23 cases from 1953-74 when this ratio was surpassed (and sometimes by quite a large margin). In fact, from 1985-2011, the number of times a region had a ratio value above 1.2 was five, zero, six and four, moving westward from the East to the West region, respectively.

Table 9 – Region ratios: teams with top TSR strength / second highest strength

| | East | Mideast | Midwest | West |
|---|---|---|---|---|
| 1985-2011 (Avg) | 1.1502 | 1.1104 | 1.1436 | 1.1198 |
| 1953-74 (Avg) | 1.1752 | 1.1152 | 1.2798 | 2.3769 |
| >1.35 | 4 | 2 | 6 | 11 |
| >1.5 | 1 | 1 | 2 | 6 |
| >2.0 | 0 | 0 | 1 | 2 |

Certainly, one could argue that some teams were excluded before 1975, by the one team per conference NCAA tournament invitation restriction. But three times the ratio was larger than 2.0: in 1958, when TSR's #7 and #29 teams were in the West region (2.2511); in 1960, when the

Midwest region had the #1 and #27 teams (2.4558); and finally, in 1970, with the #2 and #24 teams invited to the West region (2.1052). Other rows in Table 9 also illustrate this regional disparity. Perhaps strong conferences were not uniformly distributed across the country, thereby inherently making the Midwest and West regions a bracket where a team's path to the Final Four could be easier - everything else being equal.

To quantify how much easier, a methodology was described earlier in this article where each NCAA invitee can have a probability calculated regarding that team's chance of reaching the Final Four, given its own TSR strength value, and the strength values of the other teams in that region. Table 5 listed the highest such probability (36.26% in 2001), since the NCAA tournament raised the number of entrants to 64 in 1985. Therefore, given that almost 90% of all #1 seeds in each region were assigned a probability between 24 and 30%, from 1985-2011, 40% would seem to be an unusually high probability to reach the Final Four. During the tournament draw era, from 1953-1974, six times the highest TSR ranked team in the East region had a probability larger than 40%, nine times in the Mideast, 13 times in the Midwest and 17 times in the West. This appears to indicate that once again, the number of highly competitive teams was smaller in the Midwest and West regions when compared to the East and Mideast. (Several teams received first round byes, in the years 1953-74, therefore, it is reasonable to expect some increases in these projected probabilities because those teams played one less regional tournament game than those teams not receiving a bye.)

Of the 17 West region teams whose probability to reach the Final Four was larger than 40%, UCLA was that team seven times, and seven other times that team was from UCLA's conference, with Utah and Seattle being the other two teams with such a high probability value (implying that the Pacific Conference, whose name has changed over the years - most recently from the Pac-8, to the Pac-10 and now the Pac-12 - was the only major conference in the West region). Increasing the probability cut-off value to 50%, only two teams in the East, two in the Mideast, eight in the Midwest, and six in the West were in this select grouping over that 22 year period, so it appears that the Midwest and West regions had a higher likelihood of only having one strong team present though sometimes upsets prevented those teams from competing in the Final Four. (UCLA in 1967 had the highest probability value, 61.7%, and Wichita State had the second highest, 57.5%, in the 1965 Midwest region; both of these teams did reach the Final Four in those years.)

## Summary

The Tournament Selection Ratio (TSR) was designed as a quantitative tool to help determine how evenly balanced the NCAA tournament regions have been during the different eras of tournament selection processes that have been used since the early 1950s. Another metric was also created, the TSR strength value, to help determine the balance within the regions.

Even though there are always a few deserving teams which may have not been invited to compete in the NCAA Men's Basketball Championship tournament each year, this study has illustrated that the tournament selection committee has done a pretty good job since 1985 - when the transition to inviting 64 teams began - in distributing the teams into four, fairly well-balanced regions. Table 8 summarized evidence which illustrated that this has not always been the case.

Before 1979 (the first year that teams were seeded), the selection committee was bound to place teams into the template known as the tournament draw, and teams were committed to regions geographically by location and/or conference affiliation. It is clear from the results (summarized most notably in Tables 8 and 9) that the Midwest and West regions were considerably less balanced, and had fewer strong teams than the other two regions. Only one team, at most, from each conference could be invited to the NCAA tournament before 1975; the lifting of that restriction, and the elimination of the tournament draw template, has provided the selection committee with the ability to create a more balanced, and representative field of the teams that have demonstrated the worthiness of having the opportunity to compete for said championship.

Other examples were also provided illustrating that certain years, especially those in the transition years (1975-1978) where invitation rules were evolving to those that are in use today, included some regions which contained more strong teams than the others. Contests between highly ranked teams occurred prematurely, i.e., during early tournament rounds, more often during this transition period because of the tournament draw template being published before the season even began, and, because of some suspect pairing decisions made by the previous committees themselves.

# Appendix A

Of the eight, computer-based rating/ranking systems alluded to previously, the RPI and Sagarin ratings were succinctly described within the article itself. The power rating system (PW - Carroll, et al, 1988) attempts to quantitatively evaluate each team's strength of schedule (SOS) using a convergent, iterative strategy, and adds the SOS to the difference between a team's average offensive and defensive point totals. By applying this system to the season's raw scores, and to said scores after they have been altered to limit each game's margin of victory (MOV) to be at most one (P1), two more of the eight systems have now been described (at a high level).

The Rewards ranking system (Trono, 2007) essentially employs an exponential weighting system to compute a rating for each team. Losses subtract from a team's rating, and wins add to it. The highest rating of any opponent that a team has defeated contributes roughly 40% more to that team's overall, weighted win rating than the next best win, which adds roughly 40%, more than the next 'most impressive' win. This continues in decreasing order according to the defeated opponents' ratings (as determined by the power rating system mentioned above that ignores MOV, i.e., P1). Each team's final rating is essentially their weighted win average times their number of wins minus their loss penalties.

The Expected Difference (ED) and Discrete Rating System (DIS) were described in detail in the Appendix of another paper (Trono, 2010). Briefly, ED compares every actual game score against the most likely game score, given each team's offensive and defensive scoring averages for all that games that year. Team A's expected score = (A's Off. Avg. + B's Def. Avg.) / 2 points, and A's defense would be expected to yield (A's Def. Avg. + B's Off. Avg.) / 2 points. If team A wins by more points than that expected difference, between team A's and team B's expected point totals, that will increase team A's overall rating (and decrease B's as well) since the sum of all those score discrepancies (from said expected score) is eventually divided by the number of

games played that year. Better teams typically have higher ratings in this ED system, and less successful teams will receive lower ratings.

The SD system used in the TSR utilizes the DIS, where each team's ratings are adjusted after each game, if necessary. Every team has an integer rating, and if two teams play each other, their ratings can be used to predict a point spread for that game. If the game's actual score is reasonably close to that prediction, i.e., below a predetermined threshold, both team's rating will remain unchanged. However, if the prediction was not that close, then one team's rating will be incremented, and the other one will be decremented.

To only use data from that year, all teams start with an initial rating of zero. The season is played, and the final ratings become each team's initial rating on the next iteration, and all games for the season continue to modify these integer ratings until a set of season ending ratings are repeated. The previous study (Trono, 2010) indicated that using a threshold of 1 point most accurately predicted outcomes in hindsight, i.e., retrodictively. Since many teams do share the same integer ratings, these ties are broken by subtracting the square root of the score differential in each loss, almost certainly guaranteeing a unique rating value within this SD ('Stabilized' DIS) system.

Finally, the MP (Modified Percentage) ranking system rounds out the eight systems that comprise the TSR. This system only examines each team's winning percentage, and computes a modified percentage by examining who they defeated (and lost to) in an effort to update said percentages to become one that takes into consideration the opponents' winning percentages as well. It does so in an iterative fashion, in a manner similar to the power rating system; such techniques continue to reevaluate each team's ratings/percentages until they have all converged to within a certain, predetermined error tolerance (from those generated in the previous iteration).

In the MP system, each team's original percentage (pct) is set to (W + 1) / (W + L + 2). Then, a team's next pct has each opponent's previous pct added to their running pct total, for those teams they defeated, and (1 - previous pct) is subtracted for those opponents that they lost to. This total is then divided by the number of games a team has played, and all next pct values are normalized to be in the range of 0.05 to 0.95, so that each win (or loss) adds (or subtracts) at least some amount to the next pct total on each subsequent iteration. Once every game that season has been considered, the computed next pct for each team then becomes their previous pct, and the process continues until each difference, between a team's previous pct and their final, next pct value is smaller than 0.0001 in magnitude. At this time, the next pct, before normalization, is the team's rating, though the software could easily perform one final normalization operation after convergence has occurred; however, since MP is a ranking system, once the ratings have produced the requested team ordering, the actual, final computed percentage value itself is somewhat superfluous.

# Appendix B

As mentioned previously, several articles have attempted to evaluate if the selection committee has selected the most deserving teams for the NCAA tournament, or, such articles have described models which help to explain said selections quantitatively (after the fact). Each year, there are many teams who feel that they have been unjustly excluded when other teams are invited instead

of them. Rather than reiterate some of what has already been published (Reilly, 2004), the purpose of this Appendix is to briefly consider some cases, from 2011 back to 1985, in light of the TSR metric derived earlier in this paper. (When referring to the TSR below, it is implied that the ranking generated by the TSR was consulted to determine a team's rank in that ordered list.)

Virginia Commonwealth (VCU) reached the Final Four in 2011, but when they were announced as a #11 seed on Selection Sunday, there was an almost unanimous live uproar by all the sports broadcasters concerning that selection. In hindsight, given their level of performance in the tournament, VCU was certainly worthy of that invitation. VCU was #68 in the TSR, however, and there were 15 teams who were not invited with larger TSR values than VCU, though four of the eight computer systems had VCU as roughly the #60 team in 2011. (Another at-large, #11 seed in 2011 - Southern California - was TSR's #74 team.)

George Mason, who competes in the Colonial conference, along with VCU, was invited as a #11 seed as well in 2006 and they also earned their way to the Final Four. However, the fact that George Mason was ranked #35 by the TSR is some indication that they had collected a much stronger body of work that year (than VCU in 2011) and deserved their at-large bid. There were 11 teams who were not ranked as highly as George Mason (by the TSR) that also received at-large bids in 2006; two of them were Bradley (TSR #43) and the Air Force Academy (TSR #44), which were probably the last two at-large teams invited (as #13 seeds). Many journalists and sportscasters questioned inviting Bradley especially given Southwest Missouri State's exclusion - where Southwest Missouri State's RPI rank of #21 is the highest RPI rank of a team not invited to the NCAA tournament (TSR had Southwest Missouri State as #26); Bradley reached the Sweet Sixteen, thereby validating their selection somewhat.

Louisville was 29-4, and the #4 team in both polls at the end of the 2004-05 season, yet only received a #4 seed (probably because of the diminished respect of the selection committee for Conference USA, when Marquette and Cincinnati migrated to the enlarged Big East conference). Louisville's performance that season earned them the #4 spot in the TSR rankings, and since they reached the Final Four, it appears that perhaps the tournament selection committee members were a little too conservative when evaluating that team's resume.

It seems that when an invitation was extended to the 17-11 Alabama team in 2003, many people were shocked (Reilly, 2004). Yet, at #39 in the TSR, their level of play that year was sufficient to place the Crimson Tide ahead of five other at-large teams, including North Carolina (NC) State. The Wolfpack was the lowest ranked at-large team (according to the TSR), just behind #53 Boston College (BC). Reilly also thought that since BC defeated NC State on the road, and since their overall records were almost identical, he claimed that NC State's selection (over BC) was an obvious case of ACC (over Big East) bias by the selection committee. NC State did receive some votes in both final polls, whereas BC did not, so some voters did think more highly of NC State as well.

At 20-11, Syracuse was not invited in 2002, to the surprise of many because 20 wins had always been assumed to be an almost guaranteed invitation total (for teams competing in a major conference). But at #60 in the TSR, there were 10 other, higher ranked teams who were also not invited that year. Wyoming, a #11 seed, was also a somewhat suspect invitation that year (Reilly,

2004), as they were #63 in the RPI, with a record of 21-8; they were #51 in the TSR, though six 'more worthy' teams were listed above them. Alabama was 21-10 in 2001, and was also not invited, but since their TSR rank was #37, they were the most deserving team that was passed over that year. (TSR's #46 and #47 teams, Richmond and Mississippi State, were the two others that were also purportedly snubbed that year by the selection committee.)

Georgia's invitation in 2001 was somewhat of a surprise, given their 16-14 record, but their RPI rank (#27) was given significant weight by the selection committee that year; TSR ranked them as #51 - with four other teams uninvited teams positioned above them. On the opposite side of the RPI spectrum, New Mexico was #74 in 1999, yet they were invited as a #9 seed. Reilly claimed that the selection committee's composition and its member's conference affiliations were heavily responsible for this invitation. At 24-8, New Mexico was as worthy as other at-large recipients: #11 seed Evansville (23-9) who was one position below New Mexico in the TSR ranking; and #12 seed Alabama-Birmingham, who was just above New Mexico in the same ranking. (New Mexico received votes in both final polls, whereas the other two teams did not.) Using the TSR, Oregon, Rutgers and Xavier were the three most worthy, uninvited teams (for those three specific invitations) but TSR did agree with all the other selections - even #12 seed Southwest Missouri State (TSR #38) and #13 seed Oklahoma (TSR #42).

Florida State was 17-13 in 1998, and only 6-10 in the ACC, but a strong, early season showing, against highly regarded, non-conference opponents, boosted their overall resume. TSR had them as #44, and therefore the penultimate at-large team (as would have been determined by the TSR), just ahead of the last team TSR would have invited: #10 seed St. Louis. 18-13 Clemson was another ACC team with a losing conference record (7-9) that received an at-large bid that year - and a #6 seed; they were #26 in the TSR.

Going back almost another decade, the selection of Southern Mississippi State (20-11) in 1990 must have been a small surprise, since they were #72 in the TSR, with 18 more qualified teams ahead of them (who were not invited). Likewise in 1986, when Utah (20-9) must have received the last at-large bid (and a #14 seed), but at #69, with 16 teams above them in the TSR ranking, they seem like an unusual choice as well.

More recently, the 2012 tournament selection committee made only a few at-large invitations that were questioned - the most talked about choice being: why Iona and not Drexel? According to the TSR, three teams were invited that should not have been. Xavier, Colorado State and South Florida were ranked below Middle Tennessee State, Drexel and Oregon. The 46 highest ranked teams in the TSR played in the tournament which is the most ever (45 in 1986 was the previous high), and Middle Tennessee was #47, Iona was #48, Drexel was #50 and Oregon was #51. Teams #55-#59 were automatic bids, and Xavier was #60, CSU was #65 and South Florida #70 - though the latter team redeemed itself by winning their play-in game, indicating they were one of the last four teams selected, and South Florida then defeated #5 seed Temple. (Compared to Drexel's omission, there was not a lot of publically voiced dismay over Middle Tennessee being excluded from the NCAA tournament, possibly because of their conference affiliation. However, their TSR ranking indicates that they were deserving of serious consideration for an invitation, and in hindsight, they did win two games in the NIT - before losing to NIT runner-up Minnesota. Drexel also won two NIT games before their loss in that tournament as well.)

# References


Berry, Scott. 2000. “A Statistician Reads the Sports Page”, *Chance*, 13 (3), 56-61.

Breiter, D. J., and B. P. Carlin. 1997. “How to Play the Office Pools if You Must”, *Chance* 10(1), 5-11.

Brown, Mark, and Joel Sokol. 2010. “An Improved LRMC Method for NCAA Basketball Prediction,” *Journal of Quantitative Analysis in Sports*, 6 (3), Article 4.

Carroll, Bob, Pete Palmer, and Jim Thorn. 1988. *The Hidden Game of Football*. New York: Warner Press.

Coleman, Jay B., and Allen Lynch. 2001. “Identifying the NCAA Tournament ‘Dance Card’ “, *Interfaces*, Volume 31, Issue 3: 76-86.

Coleman, Jay B. and Allen Lynch. 2009. “NCAA Tournament Games: The Real Nitty-Gritty,” *Journal of Quantitative Analysis in Sports*, 5 (3), Article 8.

Coleman, Jay B., Michael J. Dumond, and Allen Lynch. 2010. “Evidence of Bias in NCAA Tournament Selection and Seeding”, *Management and Decision Economics*, Volume 31: 431-452.

Infoplease.com, Sports Almanac, http://www.infoplease.com/ipsa/A0747186.html

Jing, Yushi, and Andy Cox. 2008. “Crashing the Dance: Learning the Behavior of the NCAA Basketball Selection Committee”. www.cc.gatech.edu/~yjing/work/CS6230_project1.pdf, last accessed January 14, 2008.

NCAA Division I Men’s Basketball Championship Principles and Procedures for Establishing the Bracket, www.ncaamarchmadness2008.com/mens/pdf/Bracket_Prin_Proc_070106.pdf.

Petersen, Leo H. 1951. Sports Editor for the United Press: “The National Tournaments” column. *1951 Official NCAA Basketball Guide*: 22.

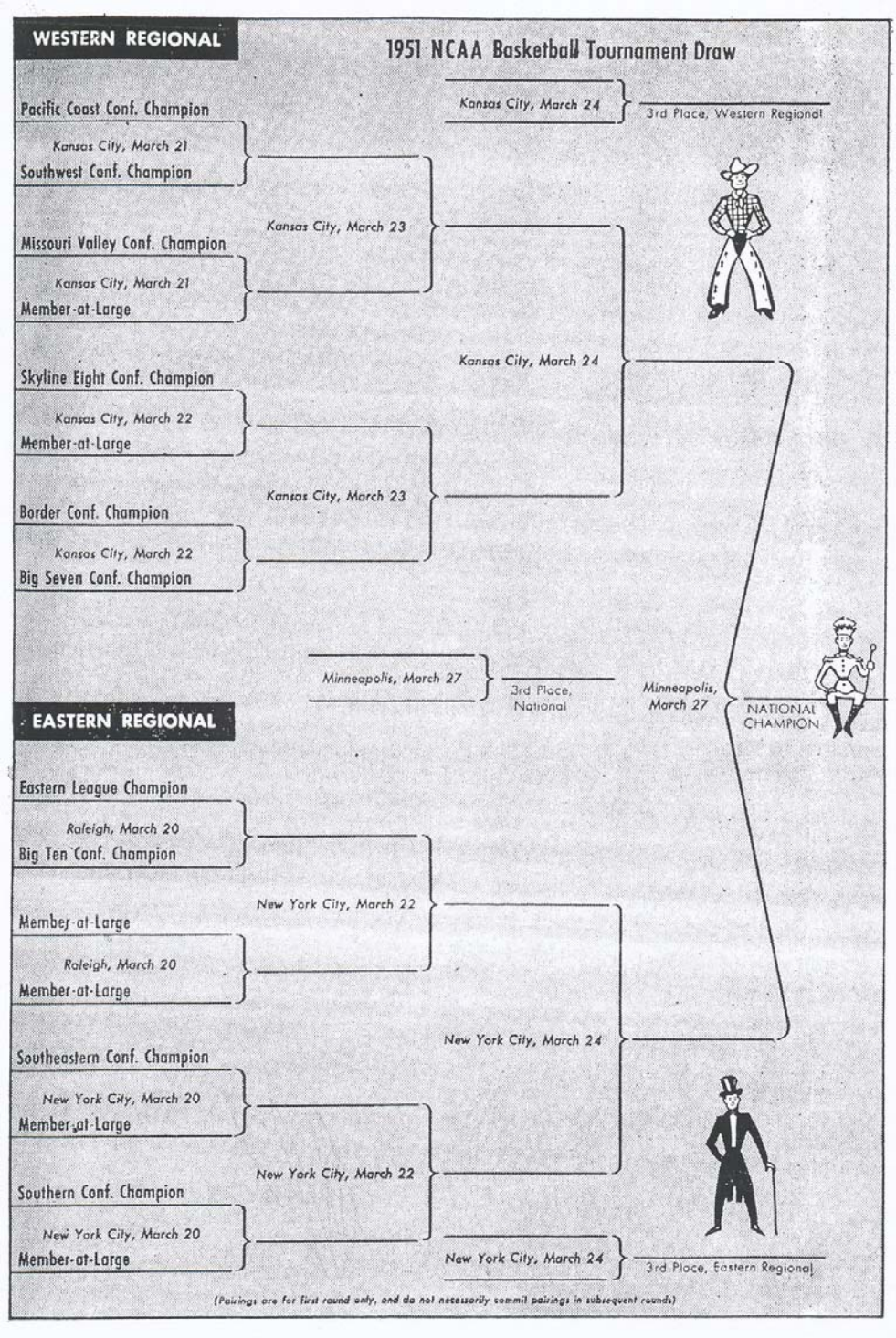
Petersen, Leo H. 1953. Sports Editor for the United Press: “The National Tournaments” column. *1953 Official NCAA Basketball Guide*: 5-6.

Reilly, P. 2003. “NCAA Tournament Bubble Watch – Statistical Approach”. Available at www.bostonsportshub.com/ncaa_bubble.htm, last accessed February 13, 2012.

Trono, John. 2007. "An Effective Nonlinear Rewards-Based Ranking System", *Journal of Quantitative Analysis in Sports*. Volume 3, Issue 2: article 3.

Trono, John. 2010. “Rating/Ranking Systems, Post-Season Bowl Games, and 'The Spread' ", *Journal of Quantitative Analysis in Sports*. Volume 6, Issue 3: article 6.

West, Brady. 2008. "A Simple and Flexible Rating Method for Predicting Success in the NCAA Basketball Tournament: Updated Results from 2007," *Journal of Quantitative Analysis in Sports*, Volume 4 (2), Article 8.




Page 23 from the *1951 Official NCAA Basketball Guide.*