International Journal of Statistics and Probability; Vol. 4, No. 3; 2015
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Tournament Seeding Efficiency and Home Court Advantage: College Basketball's National Invitation Tournament

Jeremy Phillips[1], Steven B. Caudill[2] & Franklin G. Mixon, Jr.[3]

[1] Department of Economics, Rhodes College, Memphis, TN, USA

[2] Department of Economics, Florida Atlantic University, Boca Raton, FL, USA

[3] Center for Economic Education, Columbus State University, Columbus, GA, USA

Correspondence: Franklin G. Mixon, Jr., Center for Economic Education, Columbus State University, 4225 University Avenue, Columbus, GA 31907, USA. Tel: 1-706-507-5082. E-mail: mixon_franklin@columbusstate.edu.

Received: April 11, 2015 Accepted: May 3, 2015 Online Published: July 10, 2015

doi:10.5539/ijsp.v4n3p101 URL: http://dx.doi.org/10.5539/ijsp.v4n3p101

Abstract

The relationship between seeding and outcomes in the Men's NCAA Basketball Tournament has been explored by a number of relatively recent studies. However, there has been very little published research on the Men's Postseason National Invitation Tournament, which, unlike contests in the former tournament, are typically played on a home court basis. As such, the relationship between seeding and expected outcomes may differ between the two events. Results from linear probability and probit models presented here indicate that, although the intercept is higher as expected, the marginal effect of a 1-unit seeding differential in an NCAA Tournament contest is only about 60-75 percent of its NIT contest counterpart.

Keywords: tournament seeding, contest outcomes, linear probability, probit estimation

1. Introduction and Background

Over the past several years there has been a significant amount of research on tournament seeding in various sporting events. Previous studies have investigated both the optimal number of tournament participants (e.g., Swofford, Mixon and Green, 2009) and the most efficient seeding in constructing an elimination tournament (e.g., Harville, 2003). Along these lines Schwenk (2000) provides a set of guidelines for seeded tournaments encompassing the concepts of delayed confrontation and rewarded sincerity. Delayed confrontation is the idea that the two strongest teams in the elimination tournament should not face each other until there are only two teams remaining, or that the four strongest teams should not face each other until there are only four teams remaining, and so on. Rewarded sincerity is the notion that the highest seeded team should have the least arduous path to the finals of the tournament, and that the weakest team should face the most difficult path to the finals, and so on again.

If the tournament seeding is done properly, a higher seed should defeat a lower seed. The objective of this study is to examine this issue by extending earlier work on seeding in men's college basketball's NCAA tournament (hereafter NCAAT) to the National Invitation Tournament (hereafter NIT). We do so using the regression framework of Schwertman, Schenk and Holbrook (1996) to examine the accuracy and effect of seeding on NIT game outcomes. Although the NCAAT and NIT are both postseason tournaments, the NCAAT has its choice of the best teams, leaving the NIT with the remainder and a need to rely on the assignment of a home court advantage to higher seeds to remain financially viable. In comparison with Schwertman, Schenk and Holbrook (1996) and the NCAAT, we find a higher intercept, clearly indicating the presence of the home court advantage awarded by the NIT, but also a larger increase in the probability of winning associated with an increase in seed difference. This result is somewhat surprising because our suspicion was that correctly seeding teams 65 through 96 was likely to be more difficult than seeding teams 1 through 64, hence there would be information lost in the seeding.

2. College Basketball's Postseason: A Tale of Two Tournaments

There are the two major annual post-season championship tournaments in men's college basketball – the NCAAT and the NIT. Although the tournaments are similar, with seeding playing a role in the determination of the eventual winner, there are important differences between the two. The objective of the NCAAT is to crown a national champion using, roughly, the best two-thirds of the top 100 teams in the country. The NIT is trying to maximize profits from a tournament using the bottom third of the top 100 teams. These different objectives are reflected in the set-up of each tournament.

2.1 NCAAT

The NCAAT is held in March and April with the goal of determining the national champion. A tournament selection committee chooses, approximately, the best 68 teams in the nation to participate. The 68 teams are split into four groups of 16, seeded one through 16. As the goal is the determination of a national champion, tournament games are played at neutral sites so as to not provide unfair advantages.

Fan interest is high. For example, in 2012-13 the tournament averaged 10.7 million viewers per game, with the championship game between Louisville and Michigan being watched by 23.4 million viewers.[1] This fan interest translates into generous television contracts. For example, CBS and Turner Broadcasting paid $10.8 billion for the broadcast rights for the fourteen year period from 2011 to 2024.[2]

2.2 NIT

The NIT selection process is similar to that used for the NCAAT. There are two ways to gain entrance into the NIT. First, “automatic bids” are given to the regular season champions of each NCAA Division I conference that failed to win their conference tournament nor were extended an “at-large” invitation by the NCAA Tournament Committee to participate in its championship format. Second, “at-large” bids are extended by the NIT tournament committee to other teams that are generally based on their regular season performances.[3] In 2006 the NIT began seeding the 32 participating teams, thus giving each a rank within the tournament. Today, the 32 teams are spread across four separate brackets, wherein each team is seeded 1 through 8, with the 1 seeds being the “best” teams and the 8 seeds being the “worst” teams. These seedings are then used to determine both matchups and game locations.

Like the NCAAT, the NIT format begins with the highest seeds playing the lowest seeds in each bracket, such that the first-round matchups are 1 vs. 8, 2 vs. 7, 3 vs. 6, and 4 vs. 5. The second round matchups are winner of 1 vs. 8 and winner of 4 vs. 5, and winner of 3 vs. 6 and winner of 2 vs. 7. The second round winners face off for the privilege of participating in the semifinal round in New York City, where up to and including the third round, the higher-seeded team is awarded home court advantage.[4]

Although now owned by the NCAA, the NIT is in a much different financial situation than the NCAAT due to several disadvantages. The tournament consists of the bottom third of the top 100 teams and is played at the same time as the women and men are playing for national championships. The old joke is that “NIT” stands for “No Important Teams.”[5] As a result, fan interest is relatively low – even the 2012-13 championship game between Baylor and Iowa drew only 1.3 million viewers.[6]

The lack of interest translates into less financial value. ESPN bought the rights to televise the NIT *and* 24 other NCAA championships from 2012 to 2023-24 for a relatively paltry $500 million.[7] With little fan interest and many substitutes available, operating a profitable NIT is problematic. The NIT solution is to play games, with the exception of the semifinals and finals, on the higher seed’s home court. This advantage helps increase both gate revenue and the probability of an attractive match-up for the final game.

3. Previous Literature: A Brief Review

If the seeding provides useful information, a higher seed should be more likely to defeat a lower seed. This result and other related issues have been empirically investigated for the men’s NCAAT by Schwertman, Schenk and Holbrook (1996), Boulier and Stekler (1999), Caudill and Godwin (2002), Caudill (2003), Zimmer and Kuethe (2008), and Coleman, DuMond, and Lynch (2010). The value of seedings has also been investigated for tennis by del Corral (2009) and for the World Cup (soccer) by Monks and Husch (2009). Below we examine previous research with the most bearing on the present study in more detail.

Our paper has most in common with the study of Schwertman, Schenk and Holbrook (1996), who analyze NCAAT results from 1985-1994 in order to estimate the probability that the higher- seeded team wins a given NCAAT contest. Using nonlinear models to account for asymmetric differences in teams’ strength, wherein the absolute value of the seeding difference is the independent variable, these authors find a positive and significant result for the seeding difference regressor.[8] This result is supported by probit estimation in Boulier and Stekler (1999), which compares NCAAT results to those from professional tennis, which also involves seeding differentials.

The Schwertman et al. (1996) result is also supported by Caudill and Godwin (2002), who examine NCAAT contests over the 1985-1998 period. These authors attempt to improve upon previous work by using a skewed logit model in order to account for heterogeneous skewness in order to improve the accuracy of the probability predictions.[9] Caudill and Godwin (2002) find the greater the seeding differential the higher the probability that the higher-seeded team wins a given NCAAT contest, and also that the higher the seed value of the highest seed in an NCAAT game, the greater the probability that the higher seed wins. More recently, the first of the two Caudill-Godwin findings is verified by Mixon and Withers (2005), who examine NCAAT outcomes over the 1985-2004 period using linear probability and simple logit models. They find that larger seeding differentials lead to higher probabilities that the higher-seed wins. However, the idea that these probabilities are sensitive to the seeding of the favored team falls (at the 0.165 level) just outside of the usual significance levels (Mixon and Withers, 2005).

Given the differences between the NCAAT and NIT formats mentioned above, any investigation into the relationship between seeding differentials and the probability that the higher seed wins in the NIT will involve the concept of home court advantage. Harville and Smith (1994) used data from the 1991-1992 season to measure the effect of home-court advantage in college basketball. As expected, they find a positive and significant impact from playing at home, which, in some cases, exceeds four points. Relatedly, Smith and Schwertman (1999) use NCAAT data to determine whether seed differences can predict the margin of victory and whether the NCAAT committee does a successful job in seeding tournament teams. They find that seeding differences do accurately predict margin of victory, however the vast majority of the games they analyze occur during the first round, where the seed difference is greatest and the outcome is arguably easiest to predict.

4. Data, Empirical Models and Estimation Results

The data used in this study cover the first three rounds of each NIT contest. As such, our statistical results are based on only those matchups wherein the two teams receive different seeds and a home court advantage is awarded to one of the two teams. There were 260 games contested during this period, 28 in each year from 2007-2014, and 36 in 2006 due to the additional opening round games that year. Our empirical models employ two separate dependent variables – *HighSeedWins* and *ScoreDiff*. The former of these is a dummy variable equal to 1 if the higher-seeded team in an NIT matchup wins, and 0 otherwise. The latter is a continuous variable equal to the difference between the higher-seeded team's score and that of the lower-seeded team. In some cases (i.e., when the higher seed wins) *ScoreDiff* will be greater than zero, while in others it will be less than zero. Both of these variables are a function of *SeedDiff*, which is the absolute value of the seeding difference between the two teams in each NIT contest. In the *HighSeedWins* model, we expect that, when estimated by OLS, the intercept term will retain an estimate in excess of 0.5, with the excess indicating the home court advantage. As the seeding difference climbs, however, the positive value of the parameter attached to *SeedDiff* will augment the value of the intercept and result in higher estimated values of *HighSeedWins*.

The set of estimates reported in Table 1 come from the type of OLS model described above. In this case, *SeedDiff* is positive, as expected, retaining an estimate of 0.048, indicating that for every 1-unit value in seed difference the probability of victory for the higher seed increases by about 4.8 percentage points. This result, using robust standard errors (White, 1980), is significant at the .01 level (based on a *t*-ratio of 3.88).[10]

Table 1. Econometric Results (Dep Var = *HiSeedWin*)

Variable	OLS	Probit
constant	0.569*	0.118
	(10.52)	(0.77)
SeedDiff	0.048*	0.158*
[3.277, 2.09]	(3.88)	(3.64)
n	260	260
R^2	0.051	
χ^2		14.1*

Notes: The numbers in parentheses above are *t*-values based on robust standard errors, where * denotes the .01 level of significance. The numbers in brackets are the mean and standard deviation, respectively, for *SeedDiff*.

It is interesting to compare our results to the analysis of the NCAAT by Schwertman, Schenk and Holbrook (1996). The estimated intercept for their model is 0.535, while our estimate for the NIT is 0.569. Both exceed 0.5, but our estimate for the NIT is larger, which is not surprising as our higher seeds are given home court advantage. The excess (intercept minus 0.5) is about double for the NIT data. The seed differential estimate in that study – approximately 0.030 – is only about 63 percent of the size of that found for the NIT in this study, suggesting that seeding differentials may have a *smaller* impact on the outcome of games in the NCAAT. This is somewhat surprising, but suggests an increasing decrease in quality as for lower-ranked teams. Stated differently, the top two seeds in an NCAAT bracket are more closely matched than the top two seeds in an NIT bracket. A similar conclusion is reached when compared to the marginal probability estimate (from a logistic regression) of about 0.035 found in Mixon and Withers (2005), which also analyzes the impact of seeding differentials using NCAAT data.[11] Lastly, given that our model above employs a dummy variable on the left-hand side, we also employ probit estimation in order to constrain $\mathbf{x}'\beta$ to the 0-1 interval (Greene, 2003). As indicated in Table 1, the estimated intercept of this model is 0.118, while the estimate for *SeedDiff* is equal to 0.158. The latter estimate is statistically significant at the .01 level (based on a *t*-value of 3.64), and it compares as expected to that of 0.106 from Caudill and Godwin's (2002) study of NCAAT games, which do not involve home court advantages.[12]

Total probability estimates of a higher seed win in a typical NIT game, for seed differences ranging from one to seven, are provided in Table 2. The estimates from the OLS regression, which are provided in the second column, suggest that in a closely-seeded contest, the higher seed faces a probability of winning of about 0.62. That probability rises as the quality of the higher-seeded team's opponent falls, reaching a zenith of about 0.91 for a seed difference of seven.

Table 2. Comparisons of Probability Estimates

	Probability High Seed Wins		
Seed Difference	OLS	Probit	Mixon-Withers Study
7	0.905	0.902	0.767
6	0.857	0.873	0.734
5	0.809	0.838	0.697
4	0.761	0.795	0.658
3	0.713	0.745	0.617
2	0.665	0.687	0.575
1	0.617	0.623	0.531

These results are compared to estimates computed from the logit regression reported in Mixon and Withers (2005) based on results from NCAAT games. Within the context of the NCAAT, the higher seed faces a probability of winning of only about 0.53 involving a matchup of closely-seeded teams. As before, this probability rises as the quality of the higher-seeded team's opponent falls, reaching a maximum of about 0.77 for a seed difference of seven. Again, this comparison produces the expected relationship given that the NIT, unlike the NCAAT, involves home-court advantages.

Table 2 also provides a comparison between the probit model estimates in this study and the column four (of Table 2) results from Mixon and Withers (2005). Again, the estimates using the NIT sample are all substantially larger than those computed from Mixon and Withers (2005), which analyzed NCAAT games. In this case, closely-seeded contests result in a probability that the higher-seeded team wins of about 0.62. As before, as the quality of the opponent diminishes, the probability of a higher seed emerging victorious rises, reaching a peak of about 0.90 in a contest between a 1-seed and an 8-seed.

Next, we develop another basic model to explore the relationship between seed difference and score differential between teams in each NIT contest. The score difference (i.e., *ScoreDiff*) uses the higher seeded team first, and is a function of *SeedDiff*. As in the case of *HiSeedWin*, larger values of *SeedDiff* are expected to lead to larger values of *ScoreDiff*. Results from OLS estimation of this relationship are presented in Table 3.

Table 3. Econometric Results (Dep Var = *ScoreDiff*)

Variable	OLS
constant	1.392*
	(4.66)
SeedDiff	2.007†
[3.277, 2.09]	(1.70)
n	260
R^2	0.074

Notes: The numbers in parentheses above are *t*-values based on robust standard errors, where *(†) denotes the .01(.10) level of significance. The numbers in brackets are the mean and standard deviation, respectively, for *SeedDiff*.

If the seeding were true, the intercept would be zero, indicating no advantage for two equal teams playing on a neutral court. That is not the case in the NIT, and our value of 1.392 again indicates the home court advantage. The results in Table 3 indicate that for every 1-unit of seeding difference, the higher seed wins by an estimated margin of just over two points. As such, first-round contests between 1-seeds and 8-seeds are expected to result in victory margins of almost 15 points, a total that pairs with an estimated probability of winning of about 0.9 (see Table 2).

5. Concluding Remarks

Seeding and selection decisions are often very difficult and controversial when applied to collegiate sporting events, such as the men's postseason basketball tournaments – the National Invitation Tournament and the NCAA Tournament. This study examines seeding in the NIT over most of the past decade, as has been done by other researchers regarding the NCAA Tournament. Our statistical results indicate that seeding differentials in the NIT are positively related to both the probability that the higher seed wins and the score differential, again favoring the higher seed. These findings suggest that the NIT selection committee has been efficient in its effort to seed tournament participants since 2006, when seeding of teams began.

Acknowledgements

The authors thank an anonymous referee for helpful comments. The usual caveat applies.

References

Boulier, B.L., & Stekler, H.O. (1999). Are sports seedings good predictors? An evaluation. *International Journal of Forecasting*, *15*(1), 83-91.

Caudill, S.B. (2003). Predicting discrete outcomes with the maximum score estimator: The case of the NCAA Men's Basketball Tournament. *International Journal of Forecasting*, *19*(2), 313-317.

Caudill, S.B., & Godwin, N.H. (2002). Heterogeneous skewness in binary choice models: Predicting outcomes in the Men's NCAA Basketball Tournament. *Journal of Applied Statistics*, *29*(7), 991-1,001. http://www.tandfonline.com/doi/abs/10.1080/02664760220000006694

Coleman, J.B., DuMond, J.M., & Lynch, A.K. (2010). Evidence of bias in NCAA Tournament selection and seeding. *Managerial and Decision Economics*, *31*(7), 431-452. http://onlinelibrary.wiley.com/doi/10.1002/mde.1499/abstract

del Corral, J. (2009). Competitive balance and match uncertainty in grand-slam tennis: Effects of seeding system, gender, and court surface. *Journal of Sports Economics*, *10*(6), 563-581. http://jse.sagepub.com/content/10/6/563.short

Greene, W.H. (2003). *Econometric analysis*. Upper Saddle River, NJ: Prentice Hall.

Gujarati, D. (2012). *Econometrics by example*. New York, NY: Palgrave Macmillan.

Harville, D.A. (2003). The selection or seeding of college basketball or football for postseason competition. *Journal of the American Statistical Association*, *98*(461), 17-27. http://www.tandfonline.com/doi/abs/10.1198/016214503388619058

Harville, D.A., & Smith, M.H. (1994). The home-court advantage: How large is it, and does it vary from team to team? *The American Statistician*, *48*(1), 22-28. http://www.tandfonline.com/doi/abs/10.1080/00031305.1994.10476013

Mixon, F.G., Jr., & Withers, M.C. (2005). Bracketology 101: Using sports to apply probabilistic concepts from elementary econometrics. In F.G. Mixon, Jr. (Ed.), *Shaping the learning curve: Essays on economic education* (pp. 110-122). New York, NY: iUniverse, Inc.

Monks, J., & Husch, J. (2009). The impact of seeding, home continent, and hosting on FIFA World Cup results. *Journal of Sports Economics*, *10*(4), 391-408. http://jse.sagepub.com/content/10/4/391.short

Schwenk, A.J. (2000). What is the correct way to seed a knockout tournament? *American Mathematical Monthly*, *107*(2), 140-150.

Schwertman, N.C., Schenk, K.L., & Holbrook, B.C. (1996). More probability models for the NCAA regional basketball tournaments. *The American Statistician*, *50*(1), 34-38. http://www.tandfonline.com/doi/abs/10.1080/00031305.1996.10473539

Smith, T., & Schwertman, N.C. (1999). Can the NCAA Basketball Tournament seeding be used to predict margin of victory? *The American Statistician*, *53*(2), 94-98. http://amstat.tandfonline.com/doi/abs/10.1080/00031305.1999.10474438

Swofford, J.L., Mixon, F.G., Jr., & T.G. Green (2009). Can a sub-optimal tournament be optimal when the prize can be collectively consumed? The case of college football's national championship. *Applied Economics*, *41*(25), 3,215-3,223. http://www.tandfonline.com/doi/abs/10.1080/00036840701765510

White, H. (1980). A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity. *Econometrica*, *48*(4), 817-838.

Zimmer, T., & Kuethe, T. (2008). Major conference bias and the NCAA Men's Basketball Tournament. *Economics Bulletin*, *12*(17), 1-6. http://www.accessecon.com/pubs/EB/2008/Volume12/EB-07L80004A.pdf

Notes

Note 1. See http://nesn.com/2013/04/march-madness-ratings-highest-in-19-years-with-ncaa-games-averaging-10-7-million-viewers-nielsen-reports/.

Note 2. See www.ncaa.com/news/basketball-men/2010-04-21/cbs-sports-turner-broadcasting-ncaa-reach-14-year-agreement.

Note 3. In the 2014 NIT, for example, there were 19 at-large participants and 13 automatic qualifiers.

Note 4. There are, in rare cases, exceptions to this aspect of the format. In the 2014 NIT, for example, Illinois, which was seeded second in its bracket, played at lower seeded teams' venues because its own venue was undergoing renovations at the time.

Note 5. This is not entirely true as the tournament does contain a few important teams, that is, teams with large fan bases. The problem for the NIT is insuring these "important" teams advance in the tournament.

Note 6. See www.sportsmediawatch.com/2013/04/college-basketball-tv-ratings-part-3-championship-week-ncaa-tournament-and-nit/.

Note 7. See

http://fs.ncaa.org/Docs/NCAANewsArchive/2011/december/espn%2Band%2Bncaa%2Bextend%2Brights%2Bagreement%2Bthrough%2B2023-24df30.html.

Note 8. Their nonlinear model captures any difference between two contests involving a seeding differential of two, such as those involving a 1-seed and a 3-seed and a 14-seed and 16-seed.

Note 9. Upon comparison to estimation by probit and logit, Caudill and Godwin (2002) discover that their model provides more accurate probabilities.

Note 10. According to this model, in a game between an 8-seed and a 1-seed, the 1-seed faces a probability of winning of approximately 0.905.

Note 11. Mixon and Withers' (2005) estimate of 0.035 is about 73 percent of the size of that found for the NIT in this study.

Note 12. The coefficient for SeedDiff found in this study is 1.5 times larger than that from Caudill and Godwin (2002). The probit-logit coefficient transformation described in Gujarati (2012) can be used to facilitate a comparison of our results to those from Mixon and Withers (2005).