The Effects of Alternative HOME-AWAY Sequences in a Best-of-Seven Playoff Series

G. W. BASSETT and W. J. HURLEY

In the NBA and NHL, the usual playoff format is a best-of-seven series where the stronger team (based on regular season performance) is given the benefit of four games scheduled in its home building. Typical HOME-AWAY schedules are HHAAAHH (the 2-3 format) for the NBA and HHAAHAH (the 2-2 format) for the NHL. Assuming that games are independent Bernoulli trials, we show that each team's probability of winning the series is unaffected by HOME-AWAY sequencing but that the average length of a series is affected by HOME-AWAY sequencing. For instance, if one team is stronger than the other in both buildings, the 2-3 format has a higher expected number of games than does the 2-2 format. The results follow from simple probability calculations. The sporting context makes this an interesting exercise for students of statistics.

KEY WORDS: Probability; Sports; Tournament.

1. INTRODUCTION

In professional basketball and hockey, a good showing in the regular season is rewarded with home advantage in the playoffs. That is, four of the seven games are scheduled in the stronger team's building. Given that a team is more likely to win at home, this home-field advantage leads naturally to a consideration of how the HOME-AWAY sequencing affects a team's chances of winning the series and the expected length of the series. In the NBA final championship series, the usual HOME-AWAY sequencing is HHAAAHH (the 2-3 format): one team is scheduled to play two games at home, three away, and then two back home. In the National Hockey League, the sequence of choice has been HHAAHAH (the 2-2 format).

Work on the home-field advantage includes papers by Harville (1977, 1978), Harville and Smith (1994), Hurley (1993), and Stefani (1980). Harville and Smith (1994) concluded that Division I college basketball scores over the 1991–1992 season exhibit a home-field advantage. Hurley (1993) concluded that there is a home-field advantage in major league baseball. There has been comparatively little work on the effects of alternative HOME-AWAY sequences on the outcome of a series playoff. Hurley (1993) examined the effects of alternative HOME-AWAY sequences for the World Series. However, the assumption used in that analysis—that the two teams are equal in a neutral setting—is not appropriate for the NBA and NHL playoff tournaments where, generally, one team is stronger in the neutral setting.

Assuming that each game is an independent Bernoulli trial, we have the following results:

1. The probability that a team wins the series *is the same for the 2-2 and 2-3 formats.* That is, each format gives the HOME team the same probability of winning the series. In fact, this result extends to any HOME-AWAY sequence. This latter result can be proved in two ways. One is by brute-force calculation. The other relies on a simple insight about the nature of a best-of-seven series.
2. The average length of a series *does depend* on HOME-AWAY sequencing. For instance, if the HOME team is stronger than the AWAY team (which is likely in the early rounds of playoffs), then the average length of the series is longer under the 2-3 format than under the 2-2 format.
3. The probability that there will be a seventh game in the series is not affected by HOME-AWAY sequencing.

The probability invariance result, that each team has the same probability of winning regardless of HOME-AWAY sequencing, does not follow when the assumption of independent Bernoulli trials is relaxed. We give an example with state-dependent probabilities (a team that is behind tries harder).

2. THE MODEL AND RESULTS

We suppose that one team—the HOME team—has the advantage of four games scheduled in its own building. We refer to the other team as the AWAY team. Suppose that the HOME team playing in its own building wins with probability p_H and that, in the other team's building, the HOME team's probability of winning is p_A. A natural way to begin is to assume that the games are independent Bernoulli trials. Let h_i^{23} represent the probability that the HOME team wins the series in exactly i games under the 2-3 format. Define a_i^{23} to be the probability that the AWAY team wins in exactly i games under the 2-3 format. We define h_i^{22} and a_i^{22} for the 2-2 format in a similar fashion.

A brute force calculation of the HOME team's probability of winning the 2-3 format proceeds along the following lines. For the HOME team to win in exactly four games, it must win the first two games in its own park (probability p_H^2), and then the next two in its opponent's home park (probability p_A^2). Thus,

$$h_4^{23} = p_H^2 p_A^2$$

G.W. Bassett is Professor of Economics and Finance at the Department of Economics, The University of Illinois at Chicago, 601 South Morgan Street, Chicago, IL 60607-7121. W.J. Hurley is Professor, Department of Business Administration, Royal Military College, Kingston, Ontario, College, K7K 5LO (E-mail: hurley_w@rmc.ca). The authors thank Hal Stern for comments on an earlier version of this article.

To win in exactly five games, the HOME team must win the fifth game and three of the first four. This happens with probability

$$h_5^{23} = p_A\left[2p_H^2 p_A(1-p_A) + 2p_A^2 p_H(1-p_H)\right].$$

In a similar fashion we compute h_6^{23}, h_7^{23}, and $a_4^{23}, \ldots, a_7^{23}$. The probability that the HOME team wins the series, then, is

$$h^{23} = h_4^{23} + h_5^{23} + h_6^{23} + h_7^{23}.$$

a^{23} is defined in a similar fashion. h^{22} and a^{22} are calculated in the same brute force way for the 2-2 format. This gives our first two results.

2.1 Result 1

Assuming that games are independent Bernoulli trials, the probability that the HOME team wins the series is the same for the 2-3 and 2-2 formats.

In fact, we can show that these invariance results extend to *any* HOME-AWAY sequencing.

2.2 Result 2

Assuming independent Bernoulli trials, the probability that a team wins a best-of-seven playoff series is invariant to the HOME-AWAY sequencing.

The following line of reasoning was suggested by Hal Stern. Consider a second type of series, termed a full-seven series, in which all seven games are played and the team which wins at least four games wins the series. A team which wins the full-seven series would win the associated best-of-seven series. The probability that a team wins the best-of-seven series is the same as the probability that it wins at least four games in the full-seven series:

$$h = \sum_{t=0}^{3}\left\{\binom{3}{t} p_A^t(1-p_A)^{3-t} \times \sum_{s=4-t}^{4}\binom{4}{s} p_H^s(1-p_H)^{4-s}\right\}. \quad (1)$$

The first term in the curly brackets is the probability that the HOME team wins exactly t games in the other team's building; the sum within the curly brackets is the probability that the HOME team wins at least $4-t$ games in its own building. The product of these two probabilities is the probability that HOME team wins given that it wins exactly t games in the other team's building. The probability that the HOME team wins is obtained by summing over t. Clearly this probability does not depend on the sequence of HOME-AWAY games. Hence the probability the HOME team wins is invariant to the HOME-AWAY sequence.

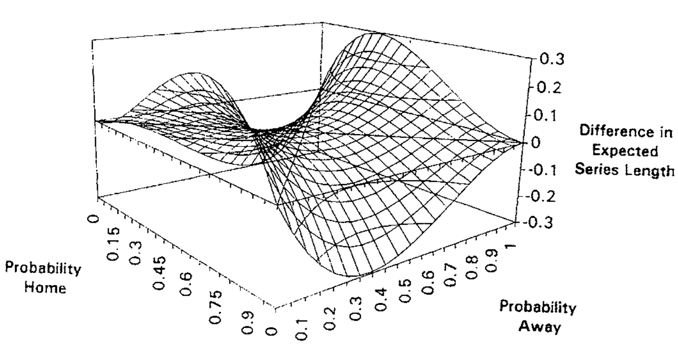

Figure 1. Graph of $\Delta(p_H; p_A)$.

2.3 Result 3

Assuming that games are independent Bernoulli trials, the probability that the series goes seven games is the same for the 2-3 and 2-2 formats.

The following table presents the probability that the HOME team wins the series and the probability the series goes seven games for various

p_H, p_A	Pr(HOME wins series)	Pr(series goes 7 games)
(.7, .6)	.813	.236
(.7, .5)	.742	.283
(.7, .4)	.661	.321

These invariance results depend critically on the assumption of independent Bernoulli trials. To see this, consider the following example with state dependent probabilities. The base assumption is $p_H = .6$ and $p_A = .5$. In contrast to independence, suppose that, if the HOME team is ahead by two games, these probabilities become $p_H = .55$ and $p_A = .45$. And if the HOME team is behind by at least two games, $p_H = .7$ and $p_A = .6$. In this situation, we now have state dependent Bernoulli trials where the team that is behind by at least two games is assumed to try harder (or, equivalently, the winning team "relaxes") and, therefore, has a better chance of winning. Under these assumptions the probability that the HOME team wins the series under the 2-3 format is .6307 and under the 2-2 format, .6288. Hence our invariance results do not extend to state-dependent Bernoulli trials.

The next set of results relate to the expected length of a series under the 2-2 and 2-3 formats. The results may be of interest to TV networks and advertisers who would prefer a long series. Let the expected length of the 2-3 series be L_{23}; let the expected length of the 2-2 series be L_{22}. The difference between these two expectations is

$$\Delta = L_{23} - L_{22} = 2(p_H + p_A - 2p_H p_A) \times (p_H - p_A)(p_H + p_A - 1), \quad (2)$$

and is derived in the Appendix. Δ has a maximum at $(p_H, p_A) = (1, 1/3)$ or $(0, 2/3)$, where $\Delta(1, 1/3) = \Delta(0, 2/3) = 8/27$. We also note the symmetry relations:

$$\Delta(p_H, p_A) = \Delta(1-p_H, 1-p_A),$$

and

$$\Delta(p_H, p_A) = -\Delta(p_A, p_H).$$

Note that the sign of Δ depends on the sign of

$$(p_H - p_A)(p_H + p_A - 1), \quad (3)$$

because $p_H + p_A - 2p_H p_A$ is always bigger than 0 if p_H

and p_A are not both 1. The first term in (3) represents the home-field bias, while the second term represents the relative strength of the teams. Relative strength is indicated by the difference in the probabilities of teams winning at home. Recall that the probability that the HOME team wins at home is p_H and the probability that the AWAY team wins at home is $1 - p_A$. This is also the same as the difference between (i) the probability that the HOME team wins when away, p_A, and (ii) the probability that the AWAY team wins when it is away, $1 - p_H$. Figure 1 is a graphical depiction of Equation (2).

2.4 Result 4

Suppose there is no home field bias. Then the expected length of the series under the 2-3 and 2-2 formats is the same. Mathematically,

$$p_H = p_A \Longrightarrow L_{23} = L_{22}. \tag{4}$$

2.5 Result 5

If a *strong* team (defined by $p_H + p_A - 1 > 0$) is more likely to win at home ($p_H > p_A$), then the expected length of the 2-3 series is longer than the 2-2 series.

The following table gives expected series lengths for various combinations of p_H and p_A:

(p_H, p_A)	*HHAAAHH*	*HHAAHAH*
(.7, .5)	5.743	5.703
(.7, .4)	5.861	5.829
(.5, .4)	5.785	5.795

3. DISCUSSION

Suppose that a typical NHL matchup in the playoffs has $p_H = .7$ and $p_A = .5$. With the 2-3 format, there would be approximately four more games per 100 series. Although this gain may appear small, there are other considerations. If there are significant costs associated with a change in venue (i.e., from Toronto to Los Angeles) such as the cost of travel, player fatigue, and so on, the 2-3 format has lower costs. Moreover, if the only off days in the series are travel days, then the 2-3 format will take a shorter calendar time to complete. Finally, the large advertising revenues from telecasts of the games means that even small increases in the expected number of games translate into large increases in revenues and profits.

4. APPENDIX: DERIVATION OF EQUATION 2

The expected length of the 2-3 format is

$$L_{23} = 4\left[h_4^{23} + a_4^{23}\right] + \cdots + 7\left[h_7^{23} + a_7^{23}\right].$$

The expression for L_{22} is similar. Taking the difference between the two, and the fact that $h_4^{23} + a_4^{23} = h_4^{22} + a_4^{22}$ and $h_7^{23} + a_7^{23} = h_7^{22} + a_7^{22}$ we get that

$$\Delta = L_{23} - L_{22} = 5\left[h_5^{23} + a_5^{23} - h_5^{22} - a_5^{22}\right] + 6\left[h_6^{23} + a_6^{23} - h_6^{22} - a_6^{22}\right]. \tag{A.1}$$

But we also have that

$$h_5^{23} + a_5^{23} + h_6^{23} + a_6^{23} = h_5^{22} + a_5^{22} + h_6^{22} + a_6^{22}, \tag{A.2a}$$

which rearranges to

$$h_6^{23} + a_6^{23} - h_6^{22} - a_6^{22} = h_5^{22} + a_5^{22} - h_5^{23} - a_5^{23}. \tag{A.2b}$$

Substituting A.2b into A.1 for $h_6^{23} + a_6^{23} - h_6^{22} - a_6^{22}$ gives

$$\Delta = h_5^{22} + a_5^{22} - h_5^{23} - a_5^{23}. \tag{A.3}$$

A simple probability computation gives

$$h_5^{23} = p_A\left[2p_H^2 p_A(1-p_A) + 2p_A^2 p_H(1-p_H)\right], \tag{A.4a}$$

and

$$h_5^{22} = p_H\left[2p_H^2 p_A(1-p_A) + 2p_A^2 p_H(1-p_H)\right]. \tag{A.4b}$$

Similar calculations give

$$a_5^{23} = (1-p_A)[2(1-p_H)^2 p_A(1-p_A) + 2(1-p_A)^2 p_H(1-p_H)] \tag{A.5a}$$

and

$$a_5^{22} = (1-p_H)[2(1-p_H)^2 p_A(1-p_A). + 2(1-p_A)^2 p_H(1-p_H)]. \tag{A.5b}$$

Substituting (A.4) and (A.5) into (A.3) and simplifying gives the required expression.

[Received March 1994. Revised September 1995.]

REFERENCES

Harville, D.A. (1977), "The Use of Linear-Model Methodology to Rate High School or College Football Teams," *Journal of the American Statistical Association*, 72, 278–289.

——— (1978), "Football Ratings and Predictions via Linear Models," in *Proceedings of the Social Statistics Section*, American Statistical Association, pp. 74–82.

Harville, D.A., and Smith, M.A. (1994), "The Home-Court Advantage: How Large Is It, and Does It Vary From Team to Team?," *The American Statistician*, 48, 22–28.

Hurley, W.J. (1993), "What Sort of Tournament Should the World Series Be?," *Chance*, 6, 31–33.

Sefani, R.T. (1980), "Improved Least Squares Football, Basketball, and Soccer Predictions," *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-10, 116–123.