“Last licks”: Do they really help?

Gary A. Simon and Jeffrey S. Simonoff[1]

Abstract

Much has been written about the home field advantage in sports. Baseball and softball are unusual games, in that the rules are explicitly different for home versus visiting teams, since by rule home teams bat second in each inning (they have “last licks”). This is generally considered to be an advantage, which seems to be contradicted by the apparent weakness of the home field advantage in baseball compared to that in other sports. In this paper we examine the effect of “last licks” on baseball and softball team success using neutral site college baseball and softball playoff games. We find little evidence of an effect in baseball, but much greater evidence in softball, related to whether a game is close late in the game. In softball games that are tied at the end of an inning, batting last seems to be disadvantageous later in the game, apparently related to the chances of the team scoring first to break the tie. By also examining games where one team was playing on its home field, we are able to say something about benefits from playing at home that are not related to “last licks.”

Key words: baseball; home advantage; softball

[1] Gary A. Simon (e-mail: gsimon@stern.nyu.edu) and Jeffrey S. Simonoff (e-mail: jsimonof@stern.nyu.edu) are Professors, Leonard N. Stern School of Business, New York University, 44 West 4th Street, New York, NY 10012. The authors would like to thank Jennifer Perrick for assistance in gathering the data analyzed here, and the editor Jim Albert and two anonymous referees for helpful comments that helped to improve the paper.

1. Introduction

The existence of the "home field advantage" is well documented, in many sports, at many skill levels, in many parts of the world. Courneya and Carron (1992, page 13) define the home field advantage to be "the consistent finding that home teams in sports competitions win over 50% of the games played under a balanced home and away schedule." The first systematic exploration of the home field advantage appears to have been by Schwartz and Barsky (1977), and its existence has been repeatedly supported since; the survey articles by Courneya and Carron (1992) and Nevill and Holder (1999) report evidence for it in college and professional baseball, high school, college, and professional football, professional hockey, high school, college, and professional basketball, professional soccer, high school track, county cricket, college softball, and college field hockey.

Courneya and Carron (1992) identified four factors that could account for home field advantage: crowd factors, familiarity with local conditions, travel factors, and effects related to rule differences for the home versus the visiting team. This paper focuses on the last of these factors, although we will also say something about other aspects of home field advantage as well. Rule difference effects have not been investigated very much in the past, no doubt in part because there are few sports where such differences exist. One prominent exception is baseball (or softball). A baseball game is separated into nine separate innings, which are further split into halves (a softball game is split into seven innings). In each half-inning, one team is in the field, while the other team bats, having the opportunity to score. If the score is tied after nine innings in baseball (or after seven innings in softball), extra innings are played. In that case, the game continues until one team has more runs than the other at the end of the inning.

Typically, the home team gets "last licks"; that is, they bat second in each of the innings. This corresponds to rule 4.02 of the Official Rules of Major League Baseball, which specifies the start of a game as follows: "The players of the home team shall take their defensive positions, the first batter of the visiting team shall take his position in the batter's box, the umpire shall call 'Play' and the game shall start" (Nemec, 1999, page 233). During the latter part of the 19th century and first half of the 20th century the choice of which team would bat first was made by the home team. It was customary for the home team to bat last in each inning, but this only became an official rule starting in the 1950 season, after the Professional Baseball Official Playing Rules Committee met in New York City to codify the rules, issuing them on December 21, 1949.

During the 19th century teams often preferred to bat first, so that they would have first opportunity to bat using the game ball, which was likely to be the only new ball used during the game (Nemec, 1999, pages 62-63). Home teams batted first at least occasionally as late as 1903, including a game between New York and Baltimore on April 21 (*New York Times*, 1903). There is even some evidence that the home team batted first by rule when major league baseball first came into existence with the formation of the National League in 1876, as an article in the *Washington Post* on June 24, 1936, while discussing an old-timers game being played in Boston, stated that "Rules which existed in 1876 will prevail, which means, among other things … the home team bats first" (United Press, 1936, page 19).

Cricket is also a sport with two distinct phases of the game (where one or the other team is at bat), but in cricket the choice of who bats first is based on the captain's choice after a coin toss. See Allsopp and Clarke (2004) for a discussion of effects related to which team bats first (among other things) in cricket matches. A coin toss is also the rule given for deciding which team bats first in what are commonly considered to be the first rules of baseball, which Alexander Cartwright laid out on September 23, 1845 (`open-site.org/Sports/Baseball/History/Rules/1845__The_Original_Rules_of_Baseball`), and was the rule used in the major leagues from 1878 through 1886 (Nemec, 1999, page 62).

Batting last in each inning of a baseball or softball game is viewed as an advantage, which leads to a bit of a puzzle, when it is noted that the home field advantage is generally smaller in baseball than it is in other sports (a roughly 54% winning percentage for major league baseball teams, compared to 60-70% figures for college and professional football, basketball, and hockey, for example). Could it be that, given two evenly matched teams, "last licks" really doesn't help, and perhaps even *hurts* a team?

2. Effects related to which team bats first: Previous evidence

In order to answer this question, we need to first understand why batting first or second in an inning could be advantageous to a team. Batting second (in the bottom of the inning) has long been viewed as an informational advantage because of the ability to implement offensive strategies in response to what the opposing team did in the top of the inning: using pinch hitters and pinch runners, playing for the tie or the win, and so on. An article in the *Christian Science Monitor* discussing the new rule regarding the order in which teams bat for the 1950 season noted that "Umpire Tom Connolly of the American League, who has been around for quite a while, said he never heard of a manager wanting to bat first, but that the home club has had that right since baseball was invented" (Associated Press, 1949, page 17). As a more recent example illustrating the wide acceptance of this position, columnist Tom Verducci castigated major league baseball commissioner Bud Selig in his column "Tom Verducci's View" in the September 20, 2004 issue of *Sports Illustrated*. Selig required that the Florida Marlins bat first in a game in Chicago against the Chicago Cubs, even though this was a makeup of a game originally scheduled for Miami that was played in Chicago because of Hurricane Frances (Verducci, 2004). Wright and House (1989) claimed that the strategic advantage of batting second "is so obvious as to need no verification" (page 131). Despite this, they estimated that "this accounts for about 5% of the home field advantage, certainly less than 10%" (page 132). It should be noted that these issues are most prominent in games that are close late in the game, since that is when strategic issues become more important.

This supposed advantage, however, ignores that the team in the defensive position in the bottom half of the inning has the ability to implement defensive strategies based on what they did offensively in the top of the inning, which could cancel out any offensive advantage (Courneya and Carron, 1990). Further, it could be that the strategies implemented by the last-batting team do not, in fact, improve the chances of winning, despite beliefs to the contrary. Many of these so-called "small ball" strategies, such as the sacrifice bunt and the stolen base, have been found to be relatively ineffective in leading to more runs being scored, despite their popularity (see, for example, Bennett, 1998, and the references therein). One further possible factor is that the team batting last has a potential disadvantage in games that are tied from the ninth inning on (seventh inning for softball), since the ace relievers used to secure victories (so-called “closers”) are often only brought into games when the team is ahead. The team batting last can never go ahead in games that are tied from the ninth inning on and then bring in a reliever (since they would have already won the game), so it might be that their best reliever never gets into the game. It should be noted that this use of relievers is a relatively recent phenomenon.

Theories of psychological momentum imply that scoring first should provide an advantage (Courneya, 1990), and obviously (all things being equal) the team that bats first is more likely to score first. As Courneya and Carron (1990) note, however, it could be that preventing the opposing team from scoring could also provide psychological momentum. This could be particularly true for baseball or softball, since if the team batting last scores first, that means that the team batting first has already used their three outs for that inning in the top of the inning, and thus has fewer opportunities to come back.

Given the expectation that other factors would probably have a stronger effect on home field advantage than the order in which the teams bat, teasing out the latter effect is a challenge, unless conditions can be found that remove the other effects. Courneya and Carron (1990) accomplished this by investigating win-loss records in a situation where the games were played on neutral fields. They examined data for 360 slo-pitch softball teams. Each team played 18 double-headers over the course of four months, where the teams playing alternated batting first in the two games of the double-header. The units of study were these double-headers, omitting those where the same team won both games, resulting in a total of 1120 double-headers. Note that in using this design, both games in all double-headers were either won by the team batting first, or by the team batting second. The authors examined the difference between the number of double-header victories by the team batting first and number of double-header victories by the team batting last, and found no evidence of any differences, including when looking at male players, female players, high ability players, low ability players, early season games, or late season games. The differences were assessed using chi-squared tests, and none were close to statistical significance.

Bray (2003) examined amateur baseball tournament games in Alberta, Canada. Each game either included an actual host team and visiting team (in the geographic sense), with the host team sometimes batting first and sometimes batting second, or two visiting teams (in the geographic sense). Bray found that the winning percentages of host teams (when playing visiting teams) were the same whether they batted first or second (57%), and the winning percentages of visiting teams (when playing visiting teams) when batting first or second were virtually equal to each other (and 50%). Bray did find that when dividing the games by age group there were significant "last licks" effects for two groups, but these were in opposite directions (one favoring the team batting first, one favoring the team batting second), so the results did not provide any compelling evidence for either an advantage or disadvantage for "last licks."

3. The data and models

In this work we investigate this question further by building on the work of Courneya and Carron (1990) and Bray (2003). The units of study are games played during the National Collegiate Athletic Association (NCAA) Division I (men's) baseball playoffs and (women's) softball playoffs. The baseball playoffs take the form of 16 four-team double-elimination Regional tournaments, followed by eight two-team double-elimination Super-Regional tournaments, with the eight winners moving on to the double-elimination College World Series (CWS). Prior to 1999 there were eight four-team Regionals and no Super-Regionals. Although the Regionals and Super-Regionals typically are hosted by one of the teams, the CWS is always played at Rosenblatt Stadium in Omaha, Nebraska. Prior to 2003 the women's softball playoffs were based on eight six-team double-elimination Regionals (typically hosted by one of the teams), with the eight winners going to the double-elimination College World Series at Hall of Fame Stadium in Oklahoma City, Oklahoma; since 2003 the eight Regionals have included eight teams each. Since the players on teams that make the NCAA Division I playoffs are highly skilled, these data have the advantage of removing home field advantage effects not related to the order in which teams bat for games played at a high level, something not possible in professional league play, where the home team always bats second. Note that looking at both baseball and softball games allows comparison of effects related to gender (Gayton et al., 1987), although these are naturally confounded with the differences between the two games.

In this context, the response of interest is not a continuous variable, but rather a binary one, corresponding to whether the first-batting team wins a game or the last-batting team wins it. Least squares regression models are inappropriate for such data, so all of the analyses are based on logistic regression models. For each game i, define the response y_i to be 0 if the first-batting team wins the game and 1 if the last-batting team wins. The logistic regression model assumes that y_i follows a Bernoulli distribution, and relates the underlying probability p_i to predictors $\{X_1, \ldots, X_k\}$ through the relationship

$$\log \frac{p_i}{1-p_i} = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki}.$$

This functional form is a natural one for probabilities, as it automatically accounts for the constraint that probabilities be bounded between 0 and 1. See Simonoff (2003, Chapter 9) for extensive discussion of this, and alternative, models for binary data.

Data on each baseball playoff game and each softball playoff game from 1999-2003 were gathered from the NCAA web site (`web1.ncaa.org/ncaa/mainsearch.do`). Of the 676 baseball games, there were 284 on neutral sites, while of the 508 softball games, 339 were played at neutral sites. For each game the runs scored by inning was recorded for both teams. In order to account for quality differences between teams, each team's won-loss record was also noted, as well as its place on a national poll (or if it is unranked) taken just before the start of the NCAA playoffs (top 30 in the *Collegiate Baseball Newspaper* poll for baseball, and top 25 in the *USA Today* / National Fastpitch Coaches Association poll for softball), and whether each team qualified for the postseason as a conference champion (an automatic qualifier) or as an "at-large" team.

The analyses here are based on predictors defined as the difference between the values for the two teams. The predictors used in the model are defined so that positive values favor the last-batting team, and negative values favor the first-batting team. In order to allow for teams that were unranked, we define the ranking variable through a series of indicator variables identifying if a team was ranked in the top 5, ranks 6 through

10, 11 through 15, and so on. Then, for each game, the differenced variable is defined as follows. Consider ranks 1 through 5. The variable `Rank 1-5` is then

$$\texttt{Rank 1-5} = \begin{cases} -1 & \text{if only the first-batting team is ranked 1 through 5} \\ 0 & \text{if neither team is ranked 1 through 5} \\ 0 & \text{if both teams are ranked 1 through 5} \\ +1 & \text{if only the last-batting team is ranked 1 through 5} \end{cases}$$

Variables corresponding to the other ranks are defined similarly. Thus, if the two teams are in the same ranking group (including both being unranked), all of the rank variables equal zero, if just one team is ranked then exactly one of the variables is nonzero, and if both teams are ranked, and ranked in different ranking groups, then exactly two of the variables are nonzero. The "at-large" differenced predictor is defined so that the value 1 corresponds to the first-batting team being a conference champion and the last-batting team being an at-large team, while the value -1 corresponds to the first-batting team being an at-large team and the last-batting team being a conference champion (the value 0 implies that either both teams are conference champions or both are at-large teams).

The benefit of using these differenced predictors is that any "last licks" effects are measured by the intercept term β_0, since

$$\frac{\exp(\beta_0)}{1+\exp(\beta_0)}$$

is the probability that the second-batting team wins the game, given the two teams are evenly matched (have the same winning percentage, the same ranking group, and are both either conference champions or at-large teams). If the game is being played at a neutral site, any observed effect can therefore be attributed to the benefit or cost of batting last.

4. "Last licks" effects

We start with an examination of the 284 neutral site baseball games. The fitted logistic regression for these data is summarized in Table 1 (including Wald z-statistics for significance of each coefficient). The one degree-of-freedom goodness-of-fit test proposed in Hosmer et al. (1997) does not indicate lack of fit for the model (p=.13). The slope coefficients all have the expected signs. For example, if the last-batting team is ranked in the top 5, while the first-batting team is unranked, the last-batting team has estimated 14.59 =exp(2.680) times the odds of winning the game, holding all else in the model fixed. Difference in winning percentage is not a strong predictor in the model (as evidenced by the Wald z-statistic), which is not surprising given the different levels of strength of schedule for different teams.

Interestingly, if two teams have the same winning percentage and are ranked similarly, an at-large team is favored to win the game over a conference champion, having 1.93 =exp(0.657) times the odds of winning. This is actually not surprising. Most conferences award their automatic bid to the winner of a postseason tournament, and in such cases it isn't always the best team in the conference that wins. Winning the conference tournament might give the team a bit of a boost in the rankings (and losing it might hurt the ranking of an at-large qualifier), but that doesn't change the inherent quality of the team.

There is little evidence of a "last licks" effect. The estimated intercept of .104 implies a winning probability of $\exp(.104)/[1+\exp(.104)] = .526$ for the last-batting team in a game between evenly matched opponents. This is not close to being significantly different from .5. As noted earlier, many of the arguments for a last-batting advantage are relevant for close games, but there is little evidence here of different patterns for close games. Figure 1 plots the observed proportions of games won by the last-batting team when the game is tied at the end of each inning, with a reference line at .5 superimposed. These proportions do not suggest any consistent pattern related to closeness of games, with the possible exception of a higher winning percentage when the game is tied at the end of the 8^{th} or 9^{th} innings. The corresponding figures for major league baseball for the 2001-2003 seasons (obtained using game-by-game records available at `www.retrosheet.org`), presented as the dashed line in the figure, also give no evidence of any effects related to the inning at which a game is tied. A fitted logistic regression to just the 32 neutral site baseball games tied at the end of the 8^{th} or 9^{th} innings (not shown) supports the lack of an effect, as there is little change in the estimated probability of the last-batting team winning for these games (it is now .541), and this probability is not close to being significantly different from .5.

Table 2 describes the softball data, based on 339 neutral site games. The model fits the data very well (p=.97). Once again an at-large team is favored over an evenly-matched automatic qualifier, having 2.36=exp(0.858) times the odds of winning a game holding all else in the model fixed. Over all of the games, there is no evidence for any "last licks" effect, as the estimated probability of the last-batting team winning is .500.

There is a very different pattern in close games, however. The winning percentage of the last-batting team drops dramatically for games that are tied at the end of the 3rd or later inning (Figure 2; note that the figure only goes out to 7 innings because that is the regulation length of softball games). It is possible that this figure could be misleading, since it does not account for the differences in abilities of the teams, but that does not seem to be the case. Table 3 gives the estimated probabilities of the last-batting team winning a game that has a specific scoring margin at the end of each inning, based on separate logistic regression fits for each game with that scoring margin. This table can be compared to Table 7 of Lindsey (1961), which gave corresponding probabilities for baseball games between evenly matched teams based on a model for scoring derived there.

Evenly matched teams should have equal chances of winning a game that is tied at the end of an inning, implying values of .5 across the middle row of Table 3. If there was a consistent last-batting advantage or disadvantage these values would not be .5, but they would still be expected to be reasonably stable. Stern (1994) extended Lindsey's results by modeling the scoring margin as a Brownian motion, possibly with drift. His Table 3 gives selected values for a baseball game with home field (and hence last-batting team) advantage corresponding to .34 runs, and they imply a reasonably constant win probability for games that are tied at the end of the inning of .52-.53 (this is consistent with the results given here for major league baseball in Figure 1).

In contrast, for these data, the estimated probability of the last-batting softball team winning a game tied at the end of an inning starts at .5, drops slowly for the first two innings, and then drops noticeably from the 3rd inning through the 6th inning. At least part of the reason for this seems to be the ability of each team to score first to break the tie (scoring first would certainly be expected to give a team an advantage). If the two teams are evenly matched and score independently of each other, the probability that the last-batting team scores first in an inning, given that a run is scored in that inning, is 1/3. (Define S to be a team scoring in the inning, and NS to be a team not scoring. There are three equally likely possibilities of the scoring pattern given a run is scored in the inning, {S, S}, {S, NS}, and {NS, S}, and in only one does the last-batting team score first.) For these data, the estimated probabilities of the last-batting team scoring first to break a tie (given in Table 4, and based on separate logistic regressions for games tied at the end of each inning) start higher than 1/3, and become progressively lower as the game goes on (but then return to roughly 1/3 for games tied at the end of the regulation 7 innings). It is reasonable to think that since the first-batting team breaks the tie more often than would be expected after innings 3 through 6, they would also win the game more often than would be expected (as it turns out they do).

Why might the last-batting team perform so much less effectively in close games as the games go on? One could speculate that this might be related to the tendency of last-batting teams to use "small ball" strategies in close games, which (as was noted earlier) might end up hurting the team. Given the relatively low scoring in softball games compared to baseball games (the median run total for the neutral site softball games is 5, while that for the neutral site baseball games is 12), it might be that these strategies are more attractive to coaches, but not necessarily more effective. Thus, if batting second does provide more information, apparently managers in these situations might be misusing it. It is not clear why the difficulties in scoring for the last-batting team are not apparent in games tied at the end of 7 innings (although it could simply be the small sample size), but in any event the close correspondence between breaking the tie first and ultimately winning the game seen in Tables 3 and 4 for all innings provides strong support for the belief that it is scoring first in tie games that drives much of what is going on.

5. Home field effects

The NCAA baseball and softball data also allow investigation of "true" home field effects, since in the Regional and Super-Regional rounds, one of the teams is the host team, and hence gets the benefits that come from playing at home. However, the host team is scheduled to bat first in many of the games. In this section we examine some of those effects.

Table 5 gives summaries of fitted logistic regressions for the probability of the last-batting team winning in the 186 baseball games where the true home team batted first and 196 games where the true home team batted last. The intercepts of the two models imply last-batting win probabilities of .516 and .536, respectively, assuming equally matched teams. Comparing this to the estimated probability in neutral site games, we see that the true home field effect is actually very small in NCAA playoff baseball, corresponding to only a 1-2 percentage point difference. This small an effect would probably come as a surprise to casual observers, since true home teams won 68.3% of the games where they batted first and 76.0% of the games where they batted last, but these overall proportions ignore team quality effects, which are important; the NCAA only awards host status to top teams, and the logistic regression fits show that it is this high quality that accounts for the bulk of their success as hosts. Interestingly, the estimated win probability of .536 for a true home team batting last is very similar to that for home teams in major league baseball's World Series through 1997 (Stern, 1998), and is similar to the rate noted earlier for the regular season in major league baseball as well.

The home field effect is much stronger in the softball data (Table 6). The intercepts of the two true home team models imply last-batting win probabilities of .389 (based on 70 games where the home team batted first) and .587 (based on 96 games where the home team batted last), respectively. Comparing this to the estimated 50% chance in neutral site games, we see that the true home field effect is roughly 10 percentage points in NCAA playoff softball, a major advantage, and a sizable part of the gain implied by the marginal home team win proportions of 75.7% (home team batting first) and 74.0% (home team batting last), respectively.

6. Conclusion

In this paper we have used logistic regression models to examine the existence and magnitude of "last licks" and home field effects in baseball and softball. The NCAA baseball and softball playoff data provide an ideal framework for such analyses, since the data include multiple games at neutral sites, games where a home team batted first, and games where a home team batted last. We found relatively little evidence for any effects in the baseball data, but much stronger evidence for "last licks" and home field effects in the softball data, suggesting that more study of softball data could be a fruitful research area.

A possible extension of the home field effects analysis performed here is an investigation of the so-called "home choke" effect. This concept, first suggested by Baumeister and Steinhilber (1984), hypothesizes that home teams that are close to victory perform poorly, being weighted down by "a burden of expectations, in the form of a supportive audience for a high-stakes performance" (Baumeister, 1995, page 644). Evidence for the existence of the "home choke" is decidedly mixed, with increasing supporting (see the references and discussion in Baumeister, 1995) and contradictory (see the references and discussion in Schlenker et al., 1995) evidence (see also Nevill and Holder, 1999). This question can be examined in the NCAA playoff data by studying the performance of home teams that are undefeated and are playing teams that are facing elimination, since they would presumably be subject to the home choke, if it exists. This would require more data than are available here, given the relative scarcity of such games.

References

1. Allsopp, P.E. and Clarke, S.R. (2004), "Rating teams and analyzing outcomes in one-day and test cricket," *Journal of the Royal Statistical Society, Series A*, **167**, 657-667.
2. Associated Press (1949), "Big league board revising baseball rule book for '50," *Christian Science Monitor*, November 17, 1949, 17.
3. Baumeister, R.F. (1995), "Disputing the effects of championship pressures and home audiences," *Journal of Personality and Social Psychology*, **68**, 644-648.
4. Baumeister, R.F. and Steinhilber, A. (1984), "Paradoxical effects of supportive audiences on performance under pressure: The home field disadvantages in sports championships," *Journal of Personality and Social Psychology*, **47**, 85-93.
5. Bennett, J.M. (1998), "Baseball," in *Statistics in Sport*, ed. J. Bennett, Arnold, London, 25-64.
6. Bray, S.R. (2003), "Batting last as a factor in the homefield advantage in competitive amateur baseball," *Sport Science Association of Alberta Report* (http://www.sportmedicinecouncilofalberta.ca/Baseball.pdf).
7. Courneya, K.S. (1990), "Importance of game location and scoring first in college baseball," *Perceptual and Motor Skills,* **71**, 624-626.
8. Courneya, K.S. and Carron, A.V. (1990), "Batting first versus last: Implications for home advantage," *Journal of Sport and Exercise Psychology*, **12**, 312-316.
9. Courneya, K.S. and Carron, A.V. (1992), "The home advantage in sport competitions: A literature review," *Journal of Sport and Exercise Psychology*, **14**, 13-27.
10. Gayton, W.F., Mutrie, S.A., and Hearns, J.F. (1987), "Home advantage: Does it exist in women's sports," *Perceptual and Motor Skills,* **65**, 653-654.
11. Hosmer, D.W., Hosmer, T., le Cessie, S., and Lemeshow, S. (1997), "A comparison of goodness-of-fit tests for the logistic regression model," *Statistics in Medicine*, **16**, 965-980.
12. Lindsey, G.R. (1961), "The progress of the score during a baseball game," *Journal of the American Statistical Association*, **56**, 703-728.
13. Nemec, D. (1999), *The Official Rules of Baseball*, Barnes and Noble Books, New York.

14. Nevill, A.M. and Holder, R.L. (1999), "Home advantage in sport: An overview of studies on the advantage of playing at home," *Sports Medicine*, **28**, 221-236.
15. *New York Times* (1903), "Baltimore, 2; N. Y., (A. L.) 0," April 22, 1903, 7.
16. Schlenker, B.R., Phillips, S.T., Boniecki, K.A., and Schlenker, D.R. (1995), "Championship pressures: Choking or triumphing in one's own territory?", *Journal of Personality and Social Psychology*, **68**, 632-643.
17. Schwartz, B. and Barsky, S.F. (1977), "The home advantage," *Social Forces*, **55**, 641-661.
18. Simonoff, J.S. (2003), *Analyzing Categorical Data*, Springer-Verlag, New York.
19. Stern, H.S. (1994), "A Brownian motion model for the progress of sports scores," *Journal of the American Statistical Association*, **89**, 1128-1134.
20. Stern, H.S. (1998), "Best-of-seven playoff series," *Chance*, **11(2)**, 46-49.
21. United Press (1936), "Old timers play game of baseball a la 1876," *Washington Post*, June 24, 1936, 19.
22. Verducci, T. (2004), "Tom Verducci's View," *Sports Illustrated*, September 20, 2004, 91.
23. Wright, C.R. and House, T. (1989), *The Diamond Appraised*, Simon and Schuster, New York.

Table 1. Fitted logistic regression for all neutral site baseball data (dependent variable is "second-batting team wins the game"; n=284).

Predictor	*Coefficient*	*Standard error*	z	p
Intercept	0.104	0.130	0.80	0.422
Winning pct.	1.014	1.426	0.71	0.477
Rank 1-5	2.680	0.623	4.30	<0.001
Rank 6-10	1.637	0.538	3.04	0.002
Rank 11-15	0.797	0.430	1.85	0.064
Rank 16-20	0.533	0.329	1.62	0.105
Rank 21-25	0.152	0.416	0.37	0.714
Rank 26-30	0.375	0.365	1.03	0.304
At-large	0.657	0.199	3.31	0.001

Table 2. Fitted logistic regression for all neutral site softball data (dependent variable is "second-batting team wins the game"; n=339).

Predictor	*Coefficient*	*Standard error*	z	p
Intercept	-0.002	0.126	-0.01	0.988
Winning pct.	1.745	1.190	1.47	0.142
Rank 1-5	2.048	0.442	4.63	<0.001
Rank 6-10	1.211	0.398	3.05	0.002
Rank 11-15	0.300	0.332	0.90	0.367
Rank 16-20	0.715	0.317	2.26	0.024
Rank 21-25	0.840	0.321	2.62	0.009
At-large	0.858	0.207	4.15	<0.001

Table 3. Estimated probabilities for neutral site softball data of the last-batting team winning the game, based on scoring margin at the end of each specified inning. Scoring margin is given as the number of runs the last-batting team is ahead or behind.

Scoring margin	*Inning*							
	0	*1*	*2*	*3*	*4*	*5*	*6*	*7*
≥ 2		.965	.973	.977	.983	.988	1.000	
1		.656	.739	.729	.783	.867	.874	
0	.500	.480	.460	.391	.352	.415	.335	.496
-1		.251	.233	.266	.258	.173	.163	
≤ -2		.000	.000	.059	.025	.000	.000	

Table 4. Estimated probabilities for neutral site softball data of the last-batting team scoring the next run to break a tie at the end of each specified inning. Equal probabilities of scoring for each team would imply values of .333 in each entry.

Inning							
0	*1*	*2*	*3*	*4*	*5*	*6*	*7*
.457	.420	.412	.353	.285	.348	.242	.320

Table 5. Fitted logistic regression for true home field baseball data (dependent variable is "second-batting team wins the game").

True home team bats first (n=186)

Predictor	*Coefficient*	*Standard error*	*z*	*p*
Intercept	0.063	0.276	0.23	0.818
Winning pct.	0.813	2.070	0.39	0.694
Rank 1-5	1.741	0.637	2.73	0.006
Rank 6-10	1.236	0.540	2.29	0.022
Rank 11-15	0.525	0.435	1.21	0.228
Rank 16-20	0.759	0.494	1.54	0.124
Rank 21-25	0.388	0.499	0.78	0.437
Rank 26-30	-0.121	0.566	-0.21	0.831
At-large	-0.057	0.259	-0.22	0.826

True home team bats last (n=196)

Predictor	*Coefficient*	*Standard error*	*z*	*p*
Intercept	0.143	0.279	0.51	0.609
Winning pct.	-0.648	2.376	-0.27	0.785
Rank 1-5	2.318	0.630	3.68	<0.001
Rank 6-10	1.911	0.592	3.23	0.001
Rank 11-15	1.307	0.486	2.69	0.007
Rank 16-20	1.840	0.573	3.21	0.001
Rank 21-25	1.064	0.540	1.97	0.049
Rank 26-30	0.208	0.735	0.28	0.777
At-large	1.022	0.301	3.39	0.001

Table 6. Fitted logistic regression for true home field softball data (dependent variable is "second-batting team wins the game").

True home team bats first (n=70)

Predictor	*Coefficient*	*Standard error*	*z*	*p*
Intercept	-0.453	0.464	-0.98	0.329
Winning pct.	0.144	3.861	0.04	0.970
Rank 1-5	3.156	1.419	2.23	0.026
Rank 6-10	0.282	0.981	0.29	0.774
Rank 11-15	0.582	0.983	0.59	0.554
Rank 16-20	0.807	1.128	0.72	0.474
Rank 21-25	-1.506	0.942	-1.60	0.110
At-large	0.079	0.571	0.14	0.890

True home team bats last (n=96)

Predictor	*Coefficient*	*Standard error*	*z*	*p*
Intercept	0.351	0.359	0.98	0.329
Winning pct.	7.673	3.243	2.37	0.018
Rank 1-5	0.876	0.920	0.95	0.341
Rank 6-10	0.374	0.891	0.42	0.675
Rank 11-15	0.749	0.783	0.96	0.339
Rank 16-20	1.098	0.750	1.46	0.143
Rank 21-25	1.357	0.769	1.76	0.078
At-large	0.196	0.386	0.51	0.611

Figure 1. Plot of observed proportions of games won by last-batting team in neutral site baseball games when the game is tied at the end of each inning (solid line). The corresponding figures for major league baseball during the 2001-2003 seasons are given as the dashed line.

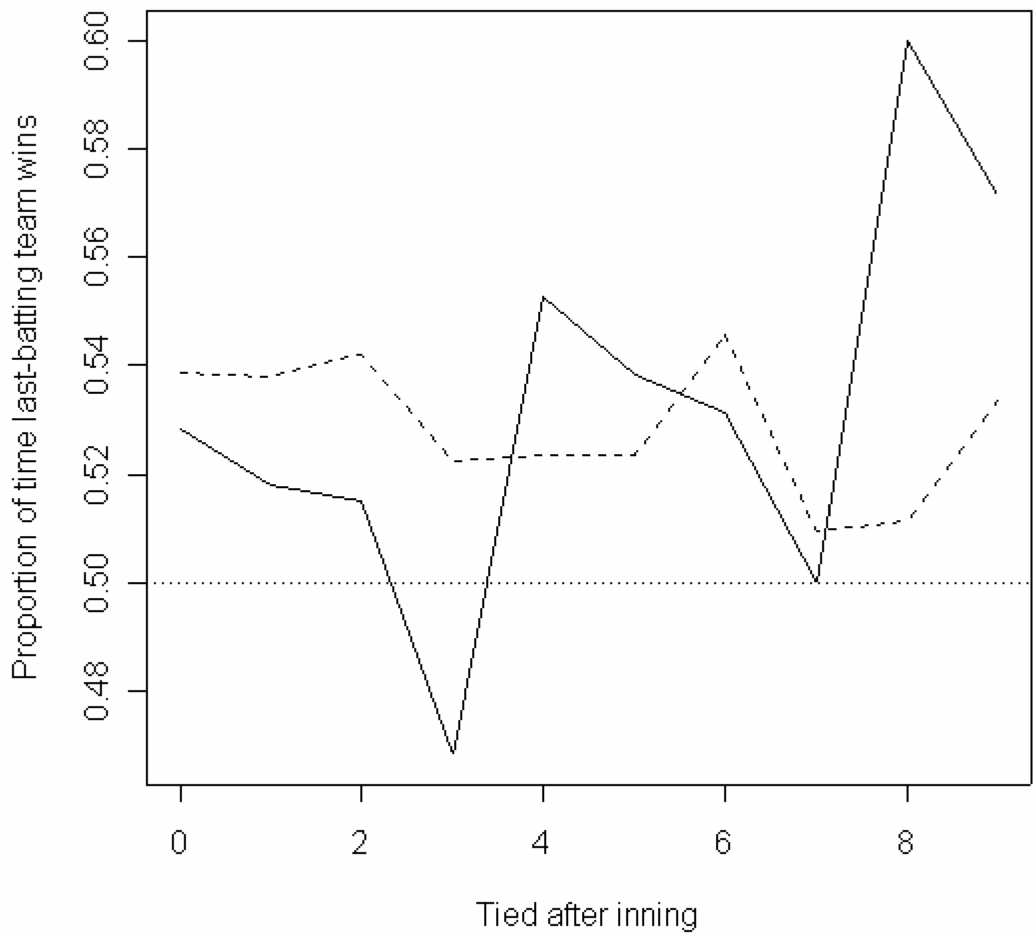

Figure 2. Plot of observed proportions of games won by last-batting team in neutral site softball games when the game is tied at the end of each inning.