By the Numbers

The Newsletter of the Statistical Analysis Committee of the Society for American Baseball Research
Volume 3, Number 4 November, 1991

Committee News

This issue of the newsletter has been somewhat delayed by my professional obligations. However, one consequence is that we have a fairly interesting set of articles for you. David and Seamus Smyth look at forecasts of pennant races made by sports journalists and compare these forecasts to a model they have developed. David Bloom provides us with an analysis of how to rank teams in a two-division league with an unbalanced schedule. Bob Davis examines the difficulties of isolating pitching and defense. Rob Wood reconsiders Hal Newhouser's career, suggesting he is more deserving of consideration for the Hall of Fame than usually thought. I look at changing performance levels over time and the difficulties that can create for analysts.

In addition, we have contributions and information from Dan Heisman, Murray Browne, and Bruce Stone.

Finally, I have included the 1992 Hall of Fame Ballot. I would encourage you to look at it and respond by casting your ballot in the Statistical Analysis Committee Hall of Fame vote.

Convention News. As you know, I was unable to attend this year's convention in New York; I have been unable to get a report on the committee meeting or a list of attendees. I do know, however, that the committee chairs suggested that each committee have an assistant chair as well as a chair. My interpretation of the justification for this is to provide for a more orderly transition when a current committee chair steps aside. I am asking, therefore, if there are any volunteers to serve as assistant chair of the Statistical Analysis Committee.

In conjunction with this, I should mention that I plan to step down as committee chair next summer. I'm hitting the wall in my ability to juggle my professional, personal, and SABR obligations. So if you are interested in being assistant chair, you should also be interested in becoming chair. The principal responsibilities of being committee chair are as follows:

1) Prepare and edit the newsletter. The two major difficulties here are extorting contributions from people and dealing with the production problems. I have been able to have Indiana University subsidize the printing and postage costs; my best guess is that these costs would total about $750 per year, which exceeds the likely budget SABR will provide (currently about $500).

2) Maintaining correspondence with committee members and keeping the mailing list up-to-date. I don't do this all that well, as many of you know.

3) Conducting the annual committee meeting at the SABR convention.

4) Working on special projects suggested or initiated by committee members. We have one of these now, which will move forward as soon as SABR figures out who its president is. I hope you all voted.

I hope that the role of the committees in SABR expands somewhat. Since I have been involved with the Statistical Analysis Committee, I have believed that the SABR research committees should have greater responsibility in helping organize and conduct the presentation part of the annual convention. The committee structure provides a useful method for organizing thematically connected research presentations at the convention. While we were able to do a little of this in Cleveland, we were the only committee to do so, and the New York organizers seemed uninterested in the idea.

If you are interested in getting more involved, let me know. I hope to have an assistant chair identified by early 1992.

Research and Publication Opportunities. Frederick Ivor-Campbell, working with the SABR Nineteenth Century Committee, writes to inform us of a publication, More Nineteenth Century Stars, which "will include biographies of all 39 nineteenth-century Hall of Famers, plus approximately 60 non-HOFers." He is editing the volume and is looking for people interested in writing bios ranging from 500 (non-HOF) to 1000 (HOF) words. If you are interested, write him at 21 Martin Street, Warren, RI 02885, or call him at 401/245-2548.

Miscellaneous. As usual, we need material. I do have three up-coming pieces already in hand. One is by David Bloom ("Lifetime Relative ERA") and two are by Bob Davis ("Does 'Save Distribution' Affect Winning Percentage?" and "An Underrated Met"). That ain't enough, folks. Let's all hit our computers, calculators, and typewriters so that you will have something to read in February.

Also, if you read an article in some other publication that's of interest, drop me a line (and enclose a copy of the article), so that other committee members can see what's happening. Talk to you in about three months.

Donald A. Coffin
Division of Business and Economics
Indiana University Northwest
3400 Broadway
Gary, IN 46408
219/980-6867

****

Forecasting Baseball Standings: Sports Journalists Versus Economists

by David J. Smyth and Seamus J. Smyth

Before the start of each baseball season, sports journalists forecast the standings of teams in each division of the American and National Leagues. This paper analyzes the accuracy of three sets of published forecasts for the nine seasons from 1982 to 1990. We compare the journalists' forecasting performance with the accuracy of forecasts generated by chance and two alternative models. The first (naive) model assumes that the standings will be the same as in the previous year. The second is a model suggested by economic theory, that the rankings of teams will depend on the relative salaries paid by the teams.

The Data. Our analysis covers the nine seasons from 1982 to 1990. We extracted predictions from the baseball preview issues of two leading sports magazines, Sport and Sports Illustrated, and from the sports section of the New York Times. The authors of these predictions forecast the standings in each division; they do not forecast win-loss percentages. We compare these rank-order forecasts with the final rank-order standings in each division.

Many studies show that forecast accuracy car improved substantially by combining forecasts.[1] see if this is the case with the journalists' forecas we constructed a new set of forecasts by combini the three sets of forecasts. This was done by firs averaging the division rankings for each season a then ordering these from one to seven (AL) or on to six (NL).

As a naive forecasting model we assume that standings will be the same as in the previous season. For the 1982 season we ignored the before-and-after-the-strike rankings and ranked teams by their winning percentages for all games played.

An Economic Model. Average salaries diffe quite widely between teams. In 1990, the averag salary of the 26 major league teams was $597,53 The Oakland As paid the highest average salary, $804,643. The lowest average salary, paid by th Baltimore Orioles, was $279,326. A plausible ut ity function for team owners consists of two com ponents. First, owner satisfaction increases with the profitability of the investment in the team. S ond, owner satisfaction increases as the team's w ning percentage, or place in the standings, rises. Empirical evidence supports the hypothesis that attendance increases with team performance.[2] Superior players cost more than inferior players.

If owners are concerned only with profit max mization, they will improve team quality up to th point that the additional revenue form an improve winning record are equal to the additional cost in salaries. If having an improved record directly increases satisfaction, owners will sacrifice some profits and improve the quality of players, and hence salaries, still further. In any event, team owners have an incentive to operate efficiently an to pay higher salaries to superior players. Thus economic theory suggests that a team's actual div sion standing should be related to its within divis relative average salaries. Thus our final model is that rankings may be predicted by the ranking of average salaries within the division.[4]

Predicting Division Winners. We first consider whether the forecasts of division winners ar more accurate than pure chance. To test this we

1. For a review and annotated bibliography, see R. T. Clemen, "Combining Forecasts: A Review and Annota Bibliography," International Journal of Forecasting, Vo 5, 1989, pp. 559-583.
2. G. W. Scully, The Business of Major League Baseball, University of Chicago Press (Chicago, IL: 1989), p. 11
3. Ibid., pp. 152-160.
4. Team salaries were obtained from various issues of the York Times and from Scully, op. cit., p. 124.

make use of the binomial distribution with probability p = 0.1429 (AL) and p = 0.1667 (NL), because the probability of correctly forecasting a division winner by pure chance, for example by drawing names out of a hat, is one-in-seven in the AL and one-in-six in the NL. There are 18 forecasts for each league, so a random forecast will be correct on average 2.57 times in the AL and three time in the NL. Do the forecasters do better than this?

Table 1: Correct Forecasts of Division Winners, 1982-1990.

Forecast	AL	NL
Sport	6**	5
Sports Illustrated	9***	3
NY Times	6**	3
Combined	5	5
Naive	4	1
Salaries	5[b]	4

***Significant at the 0.01 level.

** Significant at the 0.05 level.

* Significant at the 0.10 level.

[b] 10*** when allowance is made for the "George Steinbrener" effect.

The number of correct forecasts is shown in Table 1; instances in which the number of correct forecasts is significantly different from the number expected in random forecasts are also indicated. The journalists' prediction performance is much better than chance for the American League. However, for the NL, no set of journalists' predictions is significantly different from chance even at the 10% level. The combined journalists' forecasts are worse than the individual forecasts in the AL and are not significantly better than chance in the NL. The naive model (that this season's winner will be the same as last season's) performs poorly for both leagues. The model using relative average salaries performs better in the AL than in the NL, but is not better than chance at standard significance levels in either league. It seems that it is more difficult to *predict division winners in the NL than in the AL.* We do not know why.[5]

Next we investigate whether there is any difference in the accuracy of the forecasts between the three publications. Consider any pair of forecasters, say Sport and the New York Times. We can make pairwise comparisons, because both forecasters are trying to forecast the same event. The following procedure is used. Classify the data as follows: (i) both forecasters correctly forecast the division winner; (ii) neither forecaster correctly forecasts the division winner; (iii) Sport correctly forecasts the division winner but the New York Times does not; and (iv) the New York Times correctly forecasts the division winner but Sport does not. To test the relative forecasting ability of the two forecasters we concentrate on classifications (iii) and (iv). Table 2 gives the number of times one forecaster is correct and the other forecaster incorrect. We see that Sport had five successes when the New York Times failed and that the New York Times had three successes when Sport failed. The null hypothesis is that the proportion of successes is the same for the two forecasters. We may use a binomial distribution with p = 0.5 to see if the difference is significant. With these results, the probability that a margin of large as five to three for Sport over the New York Times is due to chance is 0.363. Thus we cannot reject the null hypothesis that the two forecasters are equally able at predicting division winners.

Similar results are given for all matched pairs in Table 2. None of the three publications outperforms the other at the 0.10 significance level. There is also no significant difference in forecasting accuracy between them and the combines forecasts series. SI, Sport, and the combined prediction series outperform the naive model (last year's winner repeats). There is no significant difference between the accuracy of the salaries model and the forecasts of the three publications or between the NYT and the naive model. If we pool the three sets of the salaries to journalists comparisons, we obtain 23 successes for the journalists against 18 for the salaries model. This difference is significant only at the 0.266 level.

5. We tested the hypothesis that there is greater parity in the NL than in the AL so that it is more difficult to forecast division winners in the NL. We examined the size of winning margins, measured by games behind, in the two leagues. The evidence did not support our hypothesis, as the winning margins were smallest in the AL East and largest in the NL East.

Table 2: Pairwise Comparisons of Forecasts of Division Winners	
Paired Comparisons[a]	Probability
SI (4) vs Sport (3)	0.500
Sport (5) vs. NYT (3)	0.363
SI (3) vs. NYT (0)	0.125
Sport (3) vs. Combined (2)	0.500
SI (2) vs. Combined (0)	0.250
Combined (2) vs. NYT (1)	0.500
Sport (6) vs. Last Season (1)	0.063*
SI (8) vs Last Season (1)	0.020**
NYT (6) vs. Last Season (2)	0.145
Combined (6) vs. Last Season (1)	0.063*
Sport (8) vs. Salaries (6)	0.395
SI (9) vs. Salaries (6)	0.304
Salaries (6) vs. NYT (6)	0.613
Combined (7) vs. Salaries	0.500
Salaries (7) vs. Last Season (3)	0.172
Journalists[b] (23) vs. Salaries (18)	0.266
The following comparisons are made excluding the New York Yankees, as described in the text.	
Salaries (12) vs. Sport (8)	0.252
Salaries (8) vs. SI (5)	0.291
Salaries (9) vs. NYT (3)	0.073*
Salaries (10) vs. Combined (5)	0.151
Salaries (12) vs. Last Season (2)	0.006***
Salaries (29) vs. Journalists[b] (16)	0.036**

[a]Successes in the paired comparison in parentheses.
[b]Journalists obtained by pooling the results for Salaries compared to Sport, Sports Illustrated, and the New York Times.

***Significant at the 0.01 level.
** Significant at the 0.05 level.
* Significant at the 0.10 level.

The salaries model assumes that owners operate efficiently. George Steinbrenner, the owner of the New York Yankees during the period covered by this study, is recognized for his high propensity to fire managers and for other activities that may reduce the productivity of his team. In every season from 1982 to 1988, the New York Yankees had the highest average salaries in the American League East, but failed to win the division title. In 1989 the Boston Red Sox had the highest average salaries; they finished second in 1989 and first in 1990. The 1990 season was the only season in which the salaries model correctly predicted the winner of the AL East. This suggests the existen of a "George Steinbrenner Effect." If we allow f such an effect and predict that the AL East will b won by the highest paid team unless the Yankees are the highest paid team, in which case the secon highest paid team is predicted to win the AL East the salary model now correctly picks the AL East winner seven times out of nine and for both AL divisions picks the winner 10 times out of 18. Th accuracy of the journalists' forecasts are unaffect by applying the same rule to them, as only twice did a publication predict that the Yankees would win the AL East; in both cases the team that was predicted to finish second also failed to win the division.

Allowing for the "George Steinbrenner Effect the salaries model predicts more winners than do the journalists. The results are shown in the seco panel of Table 2. Only the comparison with the New York Times is significant at the 0.10 level. However, if we pool the three sets of salary mod to-journalists comparisons we obtain 29 successes for the salary model against 16 for the journalists The difference is significant at the 0.036 level. T salary model also outperforms the naive model.

Predicting Division Rankings. The sports journalists forecast the complete standings of team in each division. In this section we compare thes forecasts with the final rankings. To do this we estimate the correlation coefficients between the ranks. The results are summarized in Table 3.

The first part of Table 3 gives the number of correlation coefficients that are positive, negative or zero for the AL and for the NL. If the forecas are purely random, for instance if the ordering is made by drawing the names of teams out of a hat the expected correlation coefficient is zero. Thus we can test the hypothesis that the forecaster's pe formance is better than this by using the binomia distribution with a probability of a positive correl tion coefficient of $p = 0.50$. The significance levels are shown. Eleven of the twelve sets are s nificant at at least the 0.05 level; the other is significant at the 0.071 level.

Also shown in Table 3 are the number of correlation coefficients that are significant at at least the 0.10 level, the 0.05 level, and the 0.01 level. We test the significance of these numbers using t binomial distribution with $p = 0.01$, $p = 0.05$, a $p = 0.01$ respectively. Fifteen of the eighteen results are significant at at least the 0.05 level.

Table 3: Summary Information on Correlation Coefficients Between Predicted and Actual Division Rankings, 1982-1990[a]					
	Correlation Coefficient Is Positive		Significant at		
	AL	NL	0.10	0.05	0.01
Sport	12*	14**	10***	6***	3***
SI	13**	13**	7***	4**	1
NYT	15***	15***	8***	5***	0
Combined	13**	13**	7***	4**	1
Naive	14***	15***	8***	4**	2**
Salaries	18***	13**	8***	4**	2**

[a] Number of forecasts of ranks that are positively correlated with actual finishes, or at the significance level indicated. The maximum number of trials is 18 for the AL and NL columns and 36 for the significance analysis. Significance levels are calculated using the appropriate binomial probability and are indicated as follows:
*** Significant at the 0.01 level.
** Significant at the 0.05 level.
* Significant at the 0.10 level.

This evidence strongly suggests rejecting the null hypothesis that the forecasts are no better than random forecasts. However, note that the naive and salaries models also perform very well. They always outperform the random ordering model at the 0.05 level.

We now compare the relative accuracy of the three sets of journalists forecasts with each other and with the naive and salary models. To make such comparisons we use a procedure similar to that used to compare forecasts of division winners. Again, we compare using matched pairs. Table 4 reports the number of times that one forecast generates a higher correlation coefficient than does another forecast.

There is no significant difference in the accuracy of the Sport or Sports Illustrated forecasts. However, the New York Times outperforms Sport at the 0.088 level and SI at the 0.141 level.

The next set of comparisons are between the rankings predicted by the individual publications, the combined rankings, and the naive model. There are no significant differences.

The last set of comparisons involves the salaries model. In pairwise comparisons with the individual sets of journalist predictions, the forecasts from the salary model outperforms Sport at the 0.055 level, the combined forecasts at the 0.148 level, and the naive model at the 0.088 level. The salary model had no significant advantage compared to SI or the New York Times. When we pooled the journalists' predictions, the salary model performs better by 60 to 43, which is significant at the 0.057 level. Thus we prefer the rankings of the salary model to those of the journalists.

Table 4: Pairwise Comparisons of Forecasts of Division Standings

Paired Comparisons[a]	Probability
SI (18) vs Sport (15)	0.364
NYT (22) vs. Sport (13)	0.088*
NYT (19) vs. SI (12)	0.141
Combines (19) vs. Sport (15)	0.303
SI (12) vs. Combined (11)	0.500
NYT (17) vs. Combined (12)	0.229
Sport (18) vs. Last Season (15)	0.364
Last Season (19) vs SI (14)	0.243
NYT (19) vs. Last Season (15)	0.304
Last Season(17) vs. Combined(16)	0.500
Salaries (21) vs. Sport (11)	0.055*
Salaries (19) vs. SI (16)	0.368
Salaries (20) vs. NYT (16)	0.309
Salaries (20) vs. Combined (13)	0.148
Salaries (22) vs. Last Season (13)	0.088*
Salaries (60) vs. Journalists[b] (43)	0.057*

[a]Successes in the paired comparison in parentheses.
[b]Journalists obtained by pooling the results for Salaries compared to Sport, Sports Illustrated, and the New York Times.

* Significant at the 0.10 level.

Conclusions. Our analysis of the accuracy of forecasts of division winners and of complete division rankings yields the following conclusions:

1) The forecasts published in Sport, Sports Illustrated, and the New York Times are more accurate at predicting division winners and division rankings than would be obtained randomly by drawing the names of teams out of a hat.

2) No publication had a consistently better forecasting record than did any other publication.

3) Combining the forecasts in the three publications does not result in greater forecasting accuracy.

4) The journalists' forecasts of complete division rankings are no better than those generated by assuming that this season's rankings will be the same as last season's. However, the journalists are better at forecasting the division winner than is this naive model.

5) A model that predicts performance using the relative average salaries of teams works well. The salaries model significantly outperforms the journalists at predicting division rankings. Once allowance is made for a "George Steinbrenner Effect", the salaries model is also significantly better at predicting division winners than are the journalists.

David Smyth is in the Department of Economics, Louisiana State University, Baton Rouge, LA, 70803.

****

Ranking Teams in a two-Division League

by David Bloom

General Discussion and a Rule of Thumb. Since 1969, the American and national League baseball schedules have been "divisionally balanced" as follows: Each league consists of two divisions of equal size (n teams each); each pair of teams in the same division plays c_1 games; and each pair of teams in opposite divisions plays c_2 games, where c_1. c_2 are constants (but may be different in different leagues; furthermore, c_2 was not constant in the AL in 1977 and 1978). The values of c_1 and c_2 to date have been as shown in Table 1.

Table 1: Values of n, c_1, and c_2

League	Years	n	c_1	c_2
NL	69-91	6	18	12
AL	69-76	6	18	12
AL	77-78	7	15	10,11*
AL	79-91	7	13	12

*Depending on the pair of teams.

Under such a schedule it certainly seems fair to rank teams within a division in the order of their winning percentages; it is less clear how to rank-order the league as a whole. In 1970, for example, AL interdivisional games resulted in 243 wins for East teams and 189 for West teams, indicating that the Eastern division was stronger; hence (since $c_1 > c_2$) any East team T_e was faced with a tougher schedule than a West team T_w, so that T_e could have been expected to win fewer games than T_w. Inspection of the final standings in 1970 shows

Boston	87-75
Oakland	89-73

Which team should be ranked ahead of the other in the league as a whole? (Our answer is Boston; stay tuned for the explanation of how to determine this.)

Stated in general terms, our problem is to det mine a fair rank-ordering of a divisionally balanc league, given the final standings. Complicating t problem is the fact that the mathematical model f it is not unique, because it depends on the answer the following question:
Given the respective probabilities p(i,j), p(j,k) that T_i beats T_j and that T_j beats T_k, what is th probability p(i,k) that T_i beats T_k?

Even in a game of pure chance (involving, sa a group of players who play each other in pairs w given handicaps), p(i,k) depends not just on p(i,j and p(j,k), but also in the rules of the game. In baseball we have additional factors of both skill a psychology that are impossible to pin down math ematically. These additional factors could conceivably produce three teams T_i, T_j, and T_k, such that p(i,j), p(j,k), and p(k,i) all exceed 0.50 In such circumstances, the best we can hope for i mathematical model that (a) satisfies our intuition reasonably well and (b) is not self-contradictory. condition (a) holds, we can not necessarily conclu that condition (b) also holds.[2] Various specific models for a more general ranking problem (not assuming "divisional balance") have been proposed.[3] However, I have not found in print a treatment of the special case of a two-division, di sionally balanced league; the results in this specia case are nice, and I believe them to be new.

1. (Ed. note.) An additional problem is that the observed p(i,j) (for any i, j) may not correspond to the actual, or underlying, true value of p(i,j). In any short series of trials, the observed p(i,j) will form a distribution aroun the actual p(i,j).
2. See the comment about the "uniform model" in M. Stob "A Supplement to 'A Mathematician's Guide to Popula Sports'," American Mathematical Monthly, May, 1984 pp. 277-283.
3. T. Jech, "The Ranking of Incomplete Tournaments: A Mathematician's Guide to Popular Sports," American Mathematical Monthly, April, 1983, pp. 246-266; Stot op. cit., including the bibliography.

In this article, rather than picking a specific model, I consider the class of <u>all</u> models satisfying a small set of "reasonable" assumptions (stated below). This has the advantage of generality and the disadvantage of not yielding a decision is all possible cases. However, my model assumptions have been in fact "almost" always decisive, except in seasons suffering major strike disruptions. (My method can be modified to handle a few cancellations, but not a large number or them.)

My main result is the following: To compare a team T_i in the stronger division **A** with a team T_j in the weaker division **B**, we must add to T_i's win total (and subtract from its loss total a number of the form

(1) $C + f(i)$.

The constant C is independent of the model chosen (within the class of models under consideration) and it is the same for all i, j.. In fact,

(2) $C = d(c_1 - c_2)/nc_2$,

where 2d is the excess of division **A**'s wins over division **B**'s wins. On the other hand, the term f(i) (a small error term) depends both on the model and on i and can be either positive or negative.

I will illustrate the calculation of C in the case of the comparison of Boston and Oakland cited above. We have $(n, c_1, c_2) = (6, 18, 12)$ and $2d = 243\text{-}189 = 54$. Therefore,

(2a) $C = 27(18\text{-}12)/(6*12)$
2.25

Since $87+2.25 > 89$, it follows [if we ignore the error term f(Boston)] that Boston should be ranked above Oakland. (The derivation of formulas (1) and (2) is discussed below.)

Table 2: Comparisons in Doubt When f(i) Is Included Along With C

Year	League	Comparison in Doubt
1972	NL	San Diego vs. Philadelphia
1983	NL	Montreal vs. San Diego*
1984	NL	Montreal vs. Houston
1984	NL	Montreal vs. Atlanta

*Doubt eliminated if additional assumptions are made.

Thus the problem is, in a sense, both easy and difficult. The "rule-of-thumb" constant C is easy to compute, but the error term f(i), though "small," is hard to compute even for a specific model without a sophisticated computer program. However, suitable model assumptions yield good enough <u>bounds</u> on f(i) to imply that this error term has in fact been too small to affect thr rankings in all AL and NL seasons in the 1969-1990 period, with only the exceptions shown in Table 2 (above). The value of C has been sufficient to determine the rankings in all other cases to date.

Assumptions/Restrictions and the Model. Call the 2n teams in our league $T_1, \ldots, T_{2n}$, and call the two divisions **A** and **B**, with **A** being the stronger based on interdivisional wins. Corresponding tot he partition of the league into divisions, we have a partition of the set of integers (1, ..., 2n) into subsets δ and β such that

(3) $i \in \delta \leftrightarrow T_i \in \mathbf{A}$ and $j \in \beta \leftrightarrow T_j \in \beta$

Denote the number of games won by team T_i as w_i. We assume (for simplicity) that all scheduled games are played to a decision, so we need not also consider the number of games lost. Since each team plays nc_2 games against teams in the other division, the total number of interdivisional games is n^2c_2; thus, divisions **A**, **B** win $(n^2c_2/2)+d$ and $(n^2c_2/2)\text{-}d$ respectively. (Note than $2d = \delta_{i\in\delta}w_i - \delta_{i\in\beta}w_j$ so that d is obtainable from the final standings

Suppose for the moment that the single-game level of performance of each team T_i can be quantified, that is, expressed as a continuous real-valued random variable X_i. Letting p(i,j) be the probability that $X_i > X_j$ (the probability that T_i beats T_j), suppose further that, although p(i,j) need not agree with the actual record of T_i against T_j during the season, the total number of wins (w_i, w_j) for each team is the number <u>projected</u> by the probabilities. (In statistics terminology, our estimator is "Expected Wins" = "Actual Wins", which agrees with how baseball actually ranked its teams before there were divisions!) That is, the p's satisfy the equations:

(4) $w_i = c_1\Sigma_{k\in\delta}p(i,k) + c_2\Sigma_{k\in\beta}p(i,k) \quad i \in \delta$

(5) $w_j = c_1\Sigma_{k\in\delta}p(j,k) + c_2\Sigma_{k\in\beta}p(j,k) \quad j \in \beta$

Of course, we also have

(6) $0 \le p(i,j) \le 1$ and $p(i,j) + p(j,i) = 1$

for all i, j. If we let

(7) $P_i = [1/(2n-1)]\Sigma_{k \neq i} pi,k)$

where k runs over all indices from 1 to 2n (except i), then P_i is the winning percentage predicted by the p's for T_i in the case in which $c_1 = c_2$ (i.e., if each team played each other team the same number of times--a totally balanced schedule). It thus seems natural to rank the teams according to the values of the percentages P_i: T_i is "better" than T_j (notation: T_i » T_j, or T_j « T_i if $P_i > P_j$. Clearly we can renumber the teams in such a way that

(8) $P_1 \cdot P_2 \cdot \ldots \cdot P_{2n}$

We will further assume that

(9) Whenever $p(i,j) \cdot 0.5$, then $p(i,k) \cdot p(j,k)$ for all $k = i,j$.

Assumption (9) is intuitively plausible, and (though it might not hold in reality) we can't get very far with the ranking problem without it. In particular, it can be shown that conditions (4) to (6) and (9), taken together, imply that

(10) Whenever T_i, T_j are in the same division, then T_i » T_j if and only if $w_i > w_j$.

This is in fact how baseball actually ranks teams in the same division, but it is possible for (10) to be false if (9) is not assumed [even if (4) to (6) are assumed].

By an R-Model ("R" for "rank") we shall mean a set of random variables $X_1,\ldots,X_{2n}$, together with the associated probabilities p(i,j), such that (4) to (6) and (9) hold; and, whenever discussing an R-Model, we will assume that the teams have been numbered so that (8) holds. We now assert that the **R-Model assumptions suffice to yield the quantity C + f(i) described in equations (1) and (2).** Indeed, let
$q(i,j) = p(i,j) - 0.5$ and $r(i,j,k) = q(i,j) + q(j,k) - q(i,k)$, and for $i \in \delta$ let

(11) $f(1) = (1/n)(c_1-c_2)\Sigma_{a \in \delta}\Sigma_{b \in ß} r(b,a,i)$

It then follows from (2) to (6) and (11) by an algebraic argument (whose details we omit) that, for i Ÿ Š and j Ÿ Ã,

(12) $w_i - w_j = C + f(i) + c_1\Sigma_{b \in \delta}[p(j,b)-p(i,b)]$
$+ c_2\Sigma_{a \in \delta}[p(j,a)-p(i,a)]$

By (9), the two sums on the right side of (12) are positive if T_i » T_j. Hence we conclude that T_i » is and only if $w_j - w_i < C + f(i)$, which is the desired result.

Some Specific R-Models. In view of (11), 1 would now be easy for us if we could adopt as ou model the uniform model in which all the r's are zero. Unfortunately, it is possible for the standin to be such that the uniform model is inconsistent with (4) to (6) [even without assuming (9)] and hence cannot be an R-Model. Then what sort of Model is possible? Suppose we know that

(13) No proper non-empty subset of the league wins all of its games against its compleme ary subset (the rest of the league.

Under this assumption (which holds for every sea son in the history of major league baseball), it ca be shown that there exists an R-Model for which the random variables X_i are independent and normally distributed and have a common variance. This is the "normal model", proposed by Mostell in 1951. (According to Stob, it is used at presen rating chess players.) Others, including Jech, ha proposed a different model (the "multiplicative model") which assumes that, for all i, j, k, with $i < j < k$,

(14) $[p(i,j)/p(j,i)]*[p(j,k)/pk,j)] = [p(i,k)/p(k,i)]$

This model is consistent with the R-Model assumptions, it isn't too hard to work with, and i in fact correct for some simply-defined games. B therein lies its weakness: It tends to fit simple si ations better than complicated ones, and the comp cations of baseball are innumerable. I personally prefer the "normal model", if I were forced to choose a specific model.

Another relevant remark here is that, assumin all scheduled games are played, the truth or falsit of (13) can be determined from the final standing alone (given, of course, the values of c_1 and c_2). fact, (13) is true if and only if, for every proper non-empty subset S of {1, 2, ..., 2n} (containing say s elements of δ and t elements of ß):

(15) $\Sigma_{i \in S} w_i > c_1[(s/2) + (t/2)] + c_2 st$

Even if (13) should fail, the effect is to reduce the problem to two smaller problems. These smaller problems may be more complex, since the two complementary sub-leagues may contain divisions of unequal size, but presumably, an analysis similar to that presented here would apply.

To go further and obtain useful bounds on the f(i), we need to make additional assumptions. It turns out that both the normal and multiplicative models satisfy the following three statements: whenever $i < j < k$ (so that $T_i » T_j » T_k$), then

(16a) $r(i,j,k) \cdot 0$

(16b) If i', j', and k' are indices such that $q(i',j') \cdot q(i,j)$ and $q(j'k') \cdot q(j,k)$, then $r(i',j',k') \cdot r(i,j,k)$.

(16c) $[p(i,j)/p(j,i)]*[p(j,k)/pk,j)] \leq [p(i,k)/p(k,i)]$.

[For the normal model, proofs of (16a), (16b), and (16c) are available.[4] The Jech model, as noted above, assumes <u>equality</u> in (16c), from which (16a) and (16b) can be obtained by algebra.) (16c) has the following desirable effect: If the values of p(i,j) and p(j,k) are fixed and if (16c) holds, then the value of r(i,j,k) can't exceed what its value would be in the Jech model.

We define an <u>R*-Model</u> to be an R-Model for which (16abc) hold whenever $I < j < k$. It is proved by Clarke (see fn. 4) that, whenever our model is an R*-Model with $c_1 > c_2$, then

(17) $r(i,j,k) \leq (2q^3)/[1-2q^2+(1-4q^2)^{0.5}]$

Since (11) expresses f(1) in terms of the r's, this yields bounds on the error term f(i). These and other similar results in fact suffice to show that our assertion about the ambiguous cases in Table 2 is correct if an R*-Model is assumed. If (more specifically) we assume either the normal model or the Jech (multiplicative) model), we can erase the 1983 exception. More details, including one completely worked out example, are available from the author.

For more information, contact David Bloom at 112 Division Ave., #2D, Levittown, NY, 11756.

4. (16a) and (16b) in D. Bloom, "Weighting Baseball Standings," mimeo; (16c) in L. CLarke, "Solution to Problem 5942," American Mathematical Monthly, February, 1975, pp. 186-187.

****

SHORTS I

Dan Heisman has sent me a copy of his <u>Baseball's Active Leaders Newsletter</u>, which lists the leaders in a wide range of performance categories for all active players. For example, Ricky Henderson ranks fifth in runs scored, third in walks, and first (surprise!) in stolen bases. George Brett ranks tenth (ALL-TIME) in doubles and Wade Boggs has the fifth-best (ALL-Time) career batting average (he's currently <u>ahead</u> of Ted Williams). Three of the top ten all-time stolen base leaders were active in 1991.

If you are interested in subscribing, write Dan at 1236 Paso Fino Drive, Warrington, PA 18976, or call him at 215/343-6033. The newsletter is $15 per year for four quarterly issues.

Murray Browne has a 1991 update on his research on pitching Game Scores. If you are interested in reading his analysis, which has many points of interest in it, write Murray, enclosing a self-addressed, stamped envelope, at 645 Midway, Holland, MI 49423

****

On Isolating Pitching and Defense

by Bob Davis

Runs allowed (adjusted for park effects) is obviously the best measure to describe how well a team does its job in preventing runs. However, there are actually two major components of run prevention: pitching and defense. In this paper, I will consider an index called Team Defense Rating (or TDR). TDR is an attempt partially to remove the pitching aspects of run prevention, resulting is a measure which quantifies the quality of team defense.

TDR for a given team can be intuitively described as the number of batters a pitcher on that team would have to face in order to record 27 outs, given that the pitcher allows no home runs, no walks, and no strikeouts. The logic behind this is that the defense is not responsible for any of those three possible results of a plate appearance. TDR removes those outcomes for which the pitcher is solely responsible, while retaining those outcomes which may be seen as effects of the defense.

To calculate TDR, we must first calculate the number of outs recorder by a defense. This number is given by:

(1) OUTS = AB - K - HITS + DP + CS

(where HITS is hits allowed). Note that the defense is not receiving credits for strikeouts, since these are the work of the pitcher. Since double plays are counted once in at-bats and once on their own, the team defense is receiving credit for the two outs recorded whenever a double play is turned.

Next, we need a count of the number of opportunities the defense had to record outs. This is calculated as:

(2) OPPS = AB - K - HR

Every possible at-bat outcome other than a whiff or homer means a ball must have been put in play. Whatever happens afterward is up to the defense.

From this, we can calculate the number of batters a pitcher would have to face to record a nine-inning complete game with no walks and no strikeouts as

(3) TDR = 27*(OPPS/OUTS).

Equation (3) quantifies team defense.

Qualifications. If at-bats are not available, they can be estimated by the formula 3*IP + HITS. This will automatically account for CS, since CS are credited toward innings pitched. Errors will be omitted if this estimation is used, but they cannot be simply added in because some errors allow runners to advance on the bases rather than to reach base in the first place.[1]

Technically, inside-the-park home runs should not be subtracted. However, these data are difficult to get. Double plays of the "strikeout-caught-stealing" variety should only be counted as caught steeling, as the pitcher is the one who struck out the batter. Again, it is impossible to separate these from other double plays in the standard sources. Therefore, all double plays were counted.

Ideally, only data from road games should be counted. Otherwise, park factors come into play. For example, a homer in Wrigley (not counted against the defense) might be a double or a triple Busch Stadium, making defenses in large parks lo worse than they really are.

In the calculations made in this paper, caught stealing data were unavailable, so it was omitted. Also, double-play data broken down by home/roa games were unavailable. Therefore, road double plays were approximated by total double-plays divided by two.

An Example. Here is an example of TDR, using the 1989 Pittsburgh Pirates. In road games opponents batted 2754 times, got 714 hits (including 59 homers), and struck out 392 times. The team turned 130 DP, so the road estimate is road DPs. For the 1989 Pirates:

(4) TDR = 27[(2754-392-59)/(2754-392-714+65

27*(2303/1713) = 36.30

A Pirate pitcher would need to face 36.30 hitters order for his defense to record 27 outs.

The 1989 Season. Table 1 presents the TDR calculations, from best to worst, for the NL in 1989. I have also included road ERA. These da indicate that, assuming no strikeouts, walks, or homers, an Atlanta pitcher would need to face 1. (3.2%) more batters per game than a San Francis pitcher would.

Note that there is no overwhelming relationsh between road ERA and TDR (the simple correlat between road ERA and TDR is 0.352). I believe this means that TDR is doing a good job in separating the effects of defense and pitching as they relate to runs allowed.

The Giants finished first in TFR, yet were nin in road ERA. However, the Giant pitching staff recorded only 372 road strikeouts, easily last in t league--they were leaving a lot of work for their defense to do. Furthermore, the staff allowed 59 road homers (the Phillies and Expos tied for wors in the league with 60 each). The Giants' poor performance in allowing road runs should be blamed on the pitching staff, not on the defense.

Table 1: 1989 TDR and Road ERAs

Team	TDR	Road ERA

1. (Ed. note.) If one wishes to extract data for individual pitchers from, e.g., Total Baseball, one can estimate AB as (HITS/OAV), where OAV is opponent's batting average.

SF	35.77	3.92
StL	35.87	3.47
Chi	35.96	3.31
Hou	36.01	3.71
SD	36.06	3.32
Cin	36.17	3.38
NY	36.18	3.68
Pit	36.30	4.22
Mon	36.41	3.75
LA	36.56	3.30
Phi	36.61	4.12
Atl	36.92	3.97

The Dodgers represent the opposite end of the spectrum, leading the league in road ERA but finishing a weak tenth in TDR. This discrepancy is explained by the fact that the pitchers weren't leaving that much for the fielders to do--they led the league with 513 road striekouts (no other team broke 500). The Cubs and the Dodgers had virtually identical road ERAs, but they compiled them in radically different ways. The Cubs had above-average pitching and an excellent defense, while the Dodgers had superb pitching and a poor defense.

Certainly TDR does not completely untangle pitching and defense. It is possible that a team's pitching staff could give up so many hard-hit balls that even a superior defense would compile a mediocre TDR. For example, the Phillies and the Dodgers compiled very similar TDRs, but no one would argue that the two pitching staffs gave up similar numbers of line drives. The Dodger defense has a bad TDR not because the staff was consistently hit hard, but because the defense consisted of lousy fielders.

TDR and Fielding Reputations. Does TDR agree with general opinion on which teams have the best fielders, and the best fielders at the most important positions? If it were at odds with the prevailing wisdom, either the validity of TDR or the wisdom that prevails would have to be questioned.

For the NL in 1989, TDR identifies San Francisco as the best defense, followed closely by the Cubs and Cardnials. The Giants that yeat had a regular infield of Will Clark, Robby Thompsom, Jose Uribe, and Matt Williams--all recognized as superior fielders by most experts. Brett Butler, who has an excellent reputation (and excellent fielding statistics) played center. Only the outfield flanks (Mitchell and Maldonado) would probabkly be viewed as liabilities.

The Cub infield had Mark Grace, Ryne Sandberg (probably the best defensive right-side in the last five years), and Shawon Dunston. Andre Dawson and Jerome Walton were in center and right. Only Dwight Smith in left had a reputation as a poor fielder. The Cards, with Jose Oquendo, Ozzie Smith, and Terry Pendleton in the infield, can certainly pick it. And Milt Thompson, an underrated outfielder, played flawlessly in center.

The weak teams, according to TDR, were the Dodgers, the Phillies, and the Braves. The Dodgers had an aging infield (including Eddie Murray, Alfredo Griffin, and Willie Randolph) and played both Mike Marshall and Kirk Gibson in the outfield. Only centerfielder John Shelby would be considered by most to be a good outfielder. The Phillies employed a double-play combination of Tommy Herr (who had definitely lost a step or so by this time) and Diskie Thon. The erratic Charlie Hayes was at third. As for the Braves, they had poor fielders at almost every position. Noted butcher Andres Thomas set the tone for the infield. In the outfield, 33-year-old Dale Murphy was in center, next to Lonnie Smith, who may be the most comical fielder of the decade. Also "contributing" were Ron Gant (the struggling third baseman, not the 30-30 outfielder) and the rangeless Dion James. Even Darrell Evans, in his last season, snuck into 78 games in the field. TDR definitely has this one right.

Conclusions. TDR seems to be a reasonable measure to use in quantifying team defense while attempting to remove the effects of the pitching staff. It is not a redundant measure, as it is not strongly correlated with ERA. Furthermore, for the data which have been examined so far. it has not produced any results that are wildly at odds with our qualitative judgments of fielders.

Bob Davis teaches statistics at the University of Akron; if you have comments or questions about this analysis, write him at 1504 Hunters Lake East, Cuyahoga Falls, OH 44221.

****

SHORTS II

Bruce Stone (5054 Chowen Ave. South, Minneapolis, MN 55410) has compiled games won and lost on un-earned runs, by team, for 1991. His tabulations are shwon in the following table.

Team	Wins	Losses	Pct
AL			
Oakland	10	5	.667
Minnesota	8	4	.667
Chicago	13	9	.591
New York	7	5	.583
Kansas City	8	6	.571
Toronto	9	7	.563
Boston	10	9	.526
Detroit	12	11	.522
Texas	14	13	.519
California	9	9	.500
Baltimore	7	10	.412
Seattle	8	13	.381
Milwaukee	7	12	.368
Cleveland	5	14	.263
NL			
San Diego	16	6	.727
St. Louis	13	8	.619
Philadelphia	10	8	.556
San Fran	11	9	.550
Montreal	12	12	.500
Los Angeles	11	11	.500
Pittsburgh	8	8	.500
Chicago	11	12	.478
Atlanta	10	12	.455
Houston	7	11	.389
New York	8	14	.364
Cincinnati	4	10	.286

****

Hal Newhouser: Forever Tarnished by Wartime Ball

by Rob Wood

Hal Newhouser is the only pitcher to win bac to-back Most Valuable Player Awards. He won over 200 games in a 17-year career. He retired with a fine .580 winning percentage and a very good career ERA (3.06) for his era. In Total Baseball, Pete Palmer and John Thorn rank Newhouser among the top 20 pitchers of all time But Newhouser has not received the accolades befitting a player with these achievements. For example, he has received vitrtually no considerat for the Hall of Fame. The reason for the divergence between the statistical record and the prevailing view of him as a pitcher is simple: Newhouser had his greatest seasons during World War II (29-9, 2.22 in 1944; 25-9, 1.81 in 9145).

The belief that the 1943-1945 seasons represented the lowest quality of baseball in majo league history is firmly held by baseball historian At that time, the major leagues were populated b "old men, young boys, and one-armed outfielder while the likes of DiMaggio and Williams were i military service. Further anecdotal evidence is provided by the fac tthat the lowly St. Louis Browns won their only pennant in 1944.

In this short article, I want to present evidenc on the other side of the argument. The major leagues were most assuredly weaker during the w years. However, this should not totaly discount t phenomenal seasons Newhouser put on the board

Newhouser's plight is somewhat reminiscent that of long-jumper Bob Beamon. In the 1968 Olympics, Beamon broke the world long-jump record, becoming the first ever to jump 29 feet. Although some commentators called Beamon's jump the single greatest feat in track and field history, it was also widely discounted for being accomplished in the light air of Mexico City. Th always struck me as a silly argument, since all th best long jumpers competed with Beamon in the Olympics. In addition, of course, there have bee numerous other meets held at high altitude, both before and since, and on one came close to his le --until recently, when the record was finally broken.

When I was in high school, our physics class demonstrated that Mexico City's high altitude lengthened Beamon's leap by only a few inches, whereas he obliterated the existing record by over two feet!

Newhouser's phenomenal seasons have also been disregarded on shaky grounds. Four counter arguments to the prevailing view should be made. First, in (almost) every baseball season, some pitcher(s) have outstanding seasons. It is in the statistical nature of baseball. The existence of great pitching performances does not, therefore, come as a necessary result of lowered quality of play during wartime.

Second, Newhouser's back-to-back seasons whould not go under the heading of a statistical fluke. A fluke is a pitcher with a 20-9 record and a 3.90 ERA (and 5.2 runs per game support). Prince Hal, on the other hand, was a combined 54-28 for two seasons with an ERA of 2.00.

Third, many good pitchers performed quite ordinarily during the war years. To name but a few AL hurlers, Al Smith, Bobo Newsome, Mel Harder, and Early Wynn were all unimpressive despite extensive opportunities in 1943-1945.

Fourth, purported systematic effects should be studied systematically. Instead of latching onto one connection involving one player (such as Roger Maris hitting 61 home runs in the first 162 game season), we should ask whether there was a pervasive effect. (In Maris' case, the answer is yes, but for other reasons.)

Below I present two tables of summary statistics which give evidence that far too much doubt has been cast on Newhouser's 1943-1945 records. Table 1 gives Newhouser's standard statistics for 1942 (the last season before widespread military conscription), versus his averages for 1943-1945.

Table 1: Hal Newhouser in 1942 and in 1943-1945

Performance Category	1942	1943-45
Wins	8	21
Losses	14	12
ERA	2.45	2.36
Innings	184	274
(H+W)/9IP	12.2	11.1
SO/9IP	5.0	6.0

The traditional story is that Newhouser was a lousy pitcher in 1942, then was made to look better as he faced the Pete Grays of wartime ball instead of the DiMaggios and Williamses. If you look past Newhouser's 8-14 won-loss record for the fifth place 1942 Tigers, you will see a fine 2.45 ERA (fourth best in the league).

Perhaps even more important, he led the league with over 5 strikeouts per game. Bill James and other sabermetricians have argued that strikeouts per game is one of the best indicators of future performance of a young pitcher.

Table 2 gives the same performance statistics for all other AL hurlers who pitched both in 1942 and during the war years. The criterion for inclusion in the sample for each year was 100 IP. The 1943-45 column in the average for all such seasons. The sample sizes are 33 in 1942 and 1943, 22 in 1944, and 15 in 1945.

Table 1: Other AL Pitchers in 1942 and in 1943-1945

Performance Category	1942	1943-45
Wins	10	12
Losses	11	10
ERA	3.51	3.18
Innings	181	190
(H+W)/9IP	12.2	11.7
SO/9IP	3.4	3.4

The striking feature of Table 2 is the overall similarity between 1942 and 1945-1945 for this control group. Their average won/loss record is essentially unchanged, as are innings pitched, base runners per game, and strikeouts per game. ERA did decline by 0.33, but even this improvement is not overly dramatic.

My conclusions are that wartime baseball did not automatically make an average pitcher in 1942 into a superstar of 1943-1945. Hal Newhouser was, in 1942, a very good pitcher with a bright future. His success in 1944-1945 is likely to have been genuinely achieved. After all, he led the league again in victories and ERA in 1946, the first year the veterans returned to the major leagues. I believe Newhouser deserves more consideration for inclusion in the Hall of Fame than he has received. He is, in many ways, comparable to Don Drysdale.

Rob welcomes comments on this piece. You can write him at 2101 California Street, #224, Mountain View, CA 94040.

****

Evaluating Performance in a Changing Environment

by Donald A. Coffin

I think it was Branch Rickey who once said that one mark of a pitcher with outstanding control was a strikeout-to-walk ratio (KtoW) of 2/1 or better. Even today, we can heard this referred to as a good measure of performance. The problem is that the pitching environment has changed, and what was once outstanding is now mediocre.

I compiled information on strikeouts, walks, and innings pitched for all pitchers who qualified for the ERA title in 1950 and in 1990, from Total Baseball. In 1950, if you struck out twice as many batters as you walked, you were truly unusual--only Larry Jansen (KtoW of 2.93), Preacher Roe (2.48), and Ken Raffensberger (2.17) made the cut, out of 61 ERA qualifiers--about 5%.

In 1990, 74 pitchers qualified for the ERA title, and 31 of them--42%--had KtoW ratios of 2.0 or better. To be in the top 5% required a KtoW ratio of 3.33, not 2.0 [Roger Clemens (3.87), David Cone (3.58), Danny Darwin (3.52) and Ramon Martinez (3.33)]. In fact the average KtoW ratio for all 74 ERA qualifiers in 1990 was only slightly less than 2.0 (1.96). Table 1 presents the complete distribtions.

Table 1: Distribution of Strikeout-to-Walk Ratios, 1950 and 1990

Range	Percent of ERA Qualifiers 1950	1990
2.00 +	4.9%	41.9%
1.80 - 1.99	1.6%	14.9%
1.60 - 1.79	9.8%	20.3%
1.40 - 1.59	3.3%	13.5%
1.20 - 1.39	13.1%	2.7%
1.00 - 1.19	31.1%	4.1%
0.80 - 0.99	19.7%	2.7%
< 0.80	16.4%	0.0%

Note that in 1950, more than one-third (22/61) of the ERA qualifiers walked more batters than they struck out. Only two ERA qualifiers in 1990 managed to achieve that. If Rickey was right about the importance of KtoW ratios in the early 1950s, he clearly is no longer--the environment has changed, and our performance expectations have also changed. Now, if you can't strkie out twice many batters as you walk, you aren't even averag

The change in performance has occurred both strikeouts (they are up) and walks (down). (See Table 2.) While we can debate why these change have occurred--some will argue that it is largely because of changes in hitting styles (batters dominate the outcomes), others may argue that pitching styles have more influenced--we cannot disagree that the changes have occurred.

In this respect, the debate over what has happened to the strike zone becomes even more interesting. If the strike zone has become (effectively) smaller, the rise in KtoW ratios becomes even more dramatic.

It is worthwhile, however, is to examine relative performance. We can do this by convert each pitcher's KtoW ratio into a Z-score. We do this by using the following formula

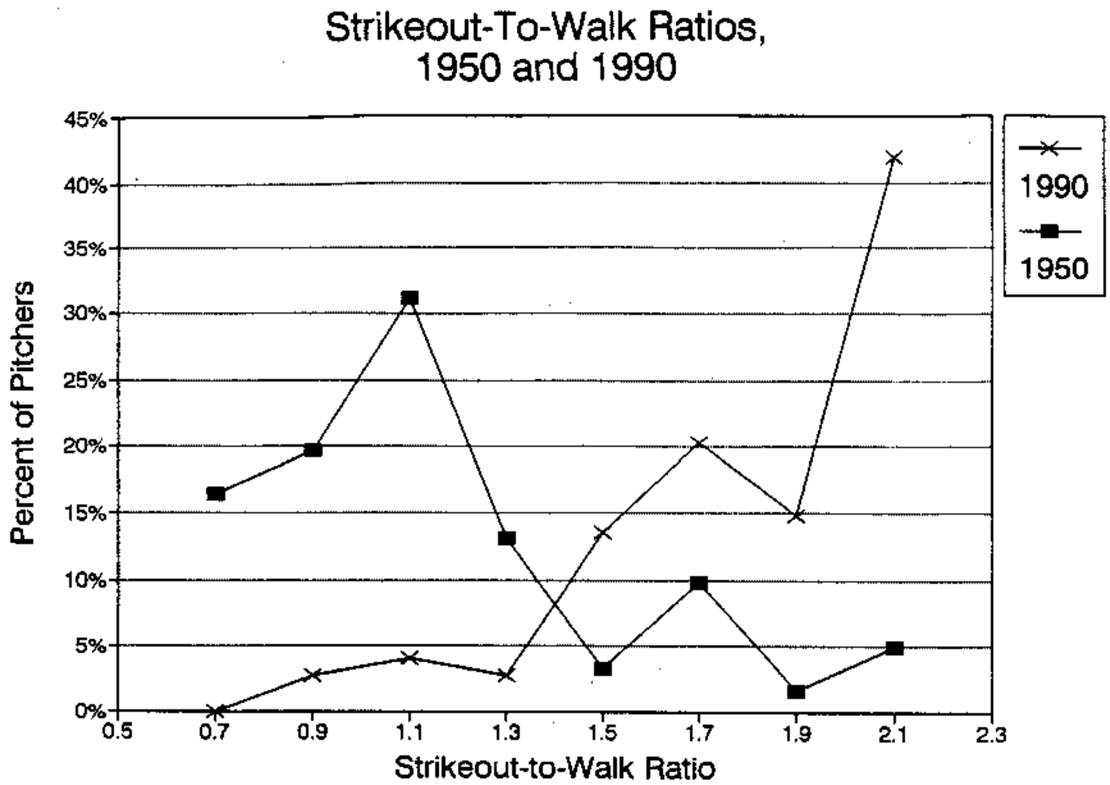
(1) $Z_i = (KtoW_i - KtoW)/\sigma_{KtoW}$

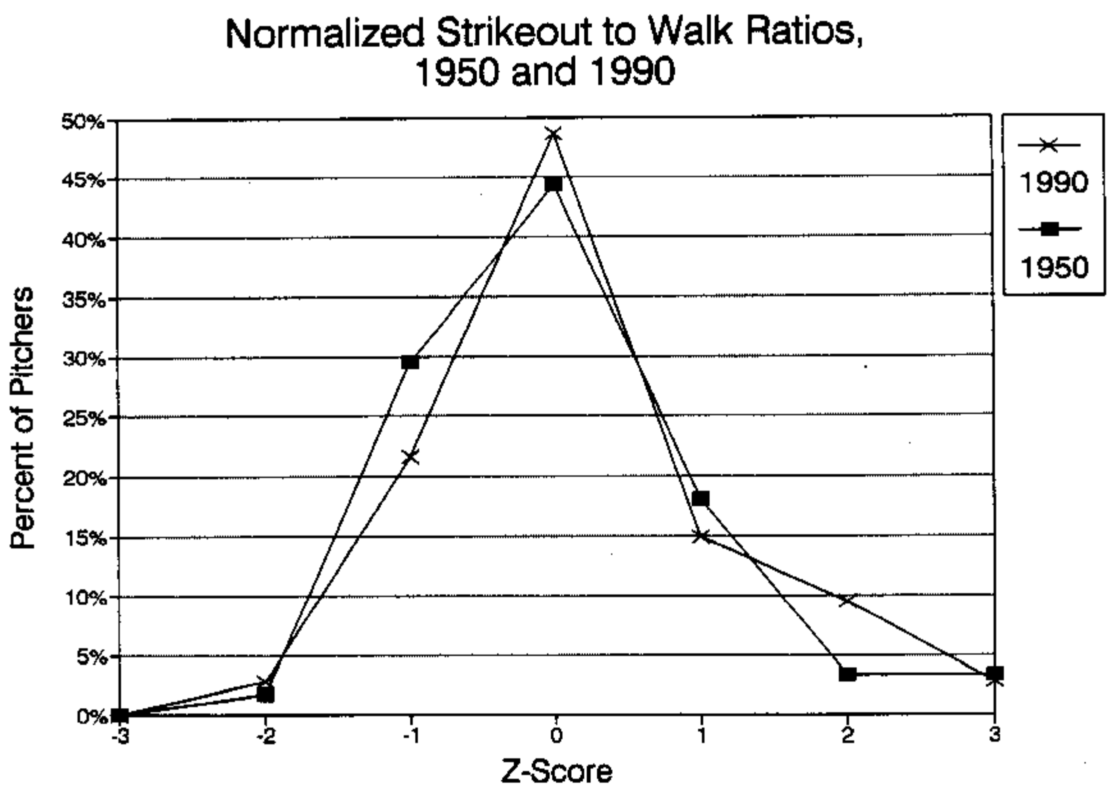
where $KtoW_i$ is an individual pitcher's KtoW rati KtoW is the ML average, and σ_{KtoW} is the standa deviation of KtoW ratios.

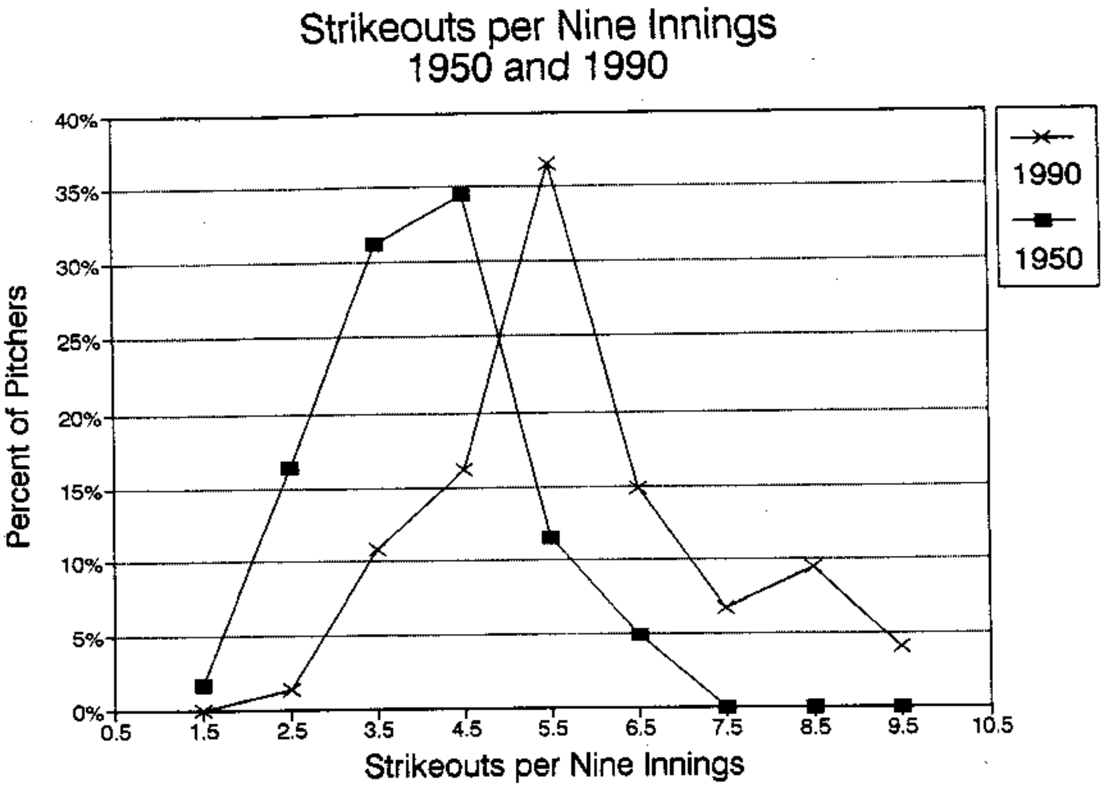
When we graph the raw KtoW ratios and the normalized KtoW ratios (Z-scores), we get the diagrams on the following page (p. 15). Note th although the distributions of unadjusted KtoW rat look very different, the normalized distributions look very much the same. In short, it is not the distribution of performance that has changed, it i the level of performance that has changed. It is s as difficult to be two standard deviations above th average as it was in 1950--but the average is dramatically higher.

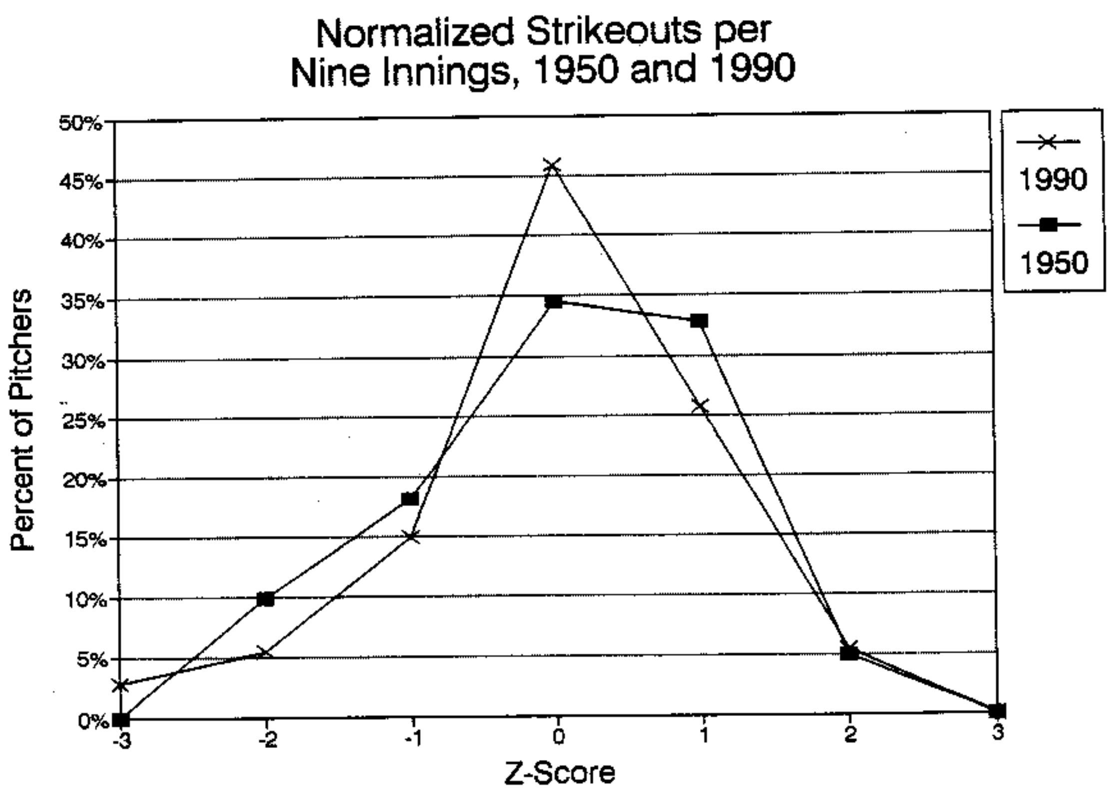
If we examine the distributions of strikeouts p nine innings (K/9IP) and walks per nine innings (W/9IP), our findings are similar. Before adjusti for changing performance levels, the distribution look different; the normalized distributions, however, look about the same. (See the diagram on pp. 16 and 17).

It is extremely important to keep changing performance levels in mind when we think about player performance over extended time periods. take one example, Hal hewhouser struck out 6.0 batters per nine innings in 1943-45 (see Rob Wood's article, above). That was outstanding fo the time. In 1990, however, it would be barely worthy of comment--too many people did that in 1990.

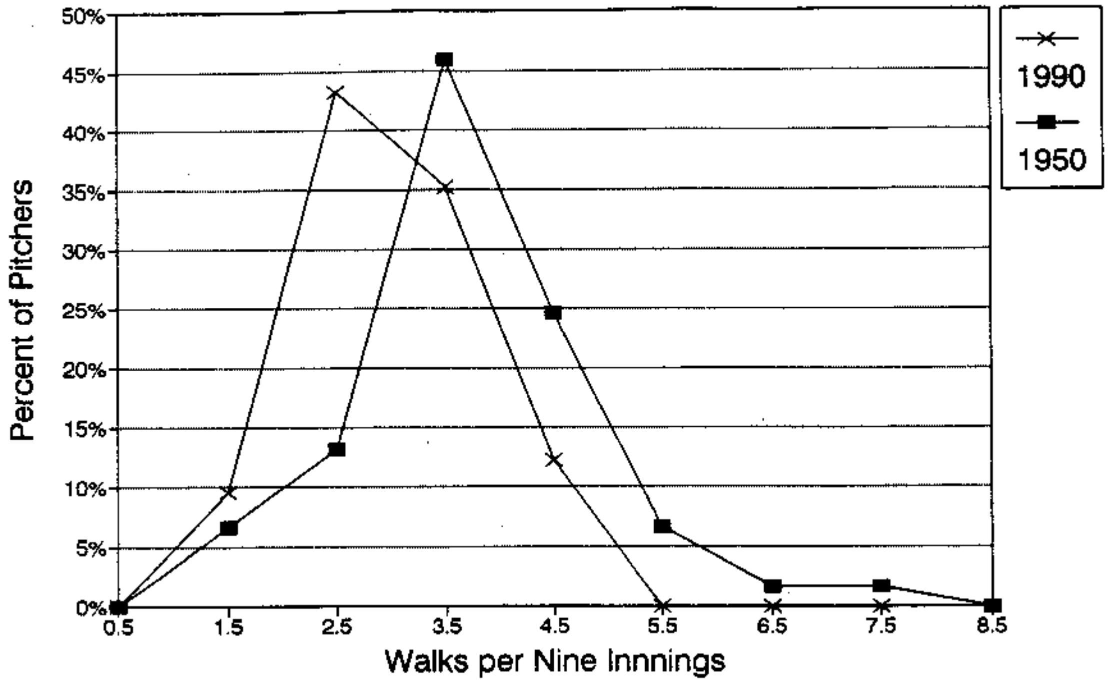

Strikeout-To-Walk Ratios,
1950 and 1990
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
Percent of Pitchers
0.5
0.7
0.9
1.1
1.3
1.5
1.7
1.9
2.1
2.3
Strikeout-to-Walk Ratio
1990
1950


Normalized Strikeout to Walk Ratios,
1950 and 1990
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
Percent of Pitchers
-3
-2
-1
0
1
2
3
Z-Score
1990
1950

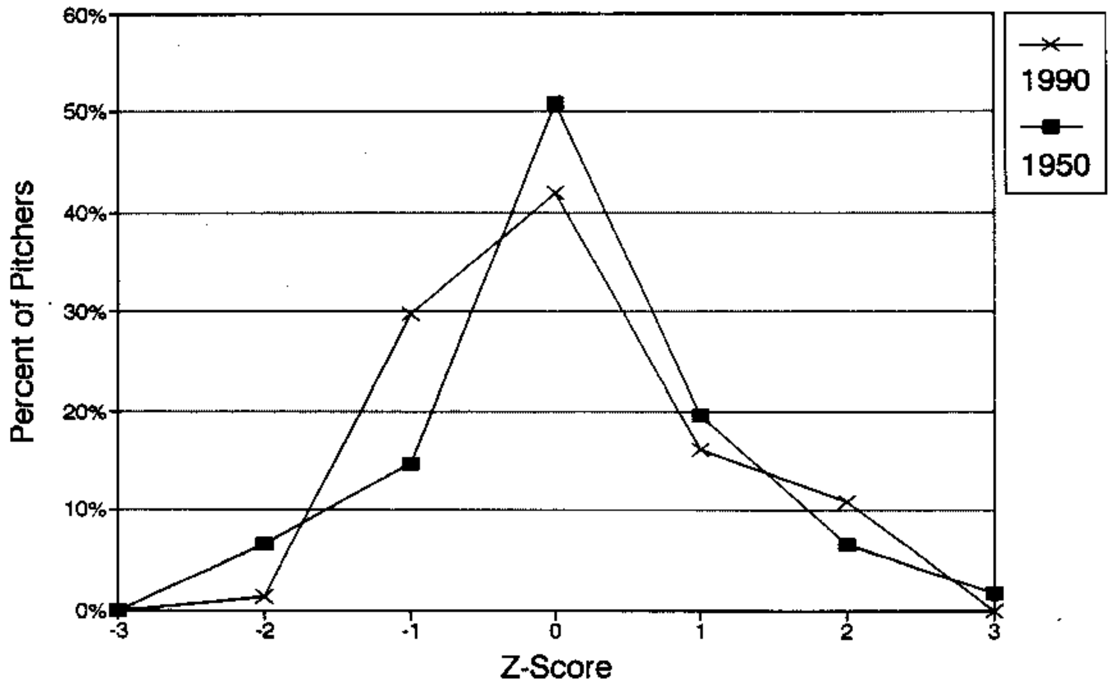

Strikeouts per Nine Innings
1950 and 1990
40%
35%
30%
25%
20%
15%
10%
5%
0%
0.5
1.5
2.5
3.5
4.5
5.5
6.5
7.5
8.5
9.5
10.5
Strikeouts per Nine Innings
Percent of Pitchers
1990
1950


Normalized Strikeouts per
Nine Innings, 1950 and 1990
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
-3
-2
-1
0
1
2
3
Z-Score
Percent of Pitchers
1990
1950

Walks per Nine Innings, 1950 and 1990



Normalized Walks per Nine Innings, 1950 and 1990



The StatComm Hall-of-Fame Ballot

The ballot for voting for the Hall of Fame candidates for 1992 was recently released. I thought it might be interesting to ask the members of the Statistical Analysis Committee to cast their ballots, using the same general rules that the Baseball Writers of America use--you can vote for up to 10 of the players named on the ballot.

To vote, write the names of the players for whom you are voting on a sheet of paper and send it to me. My address is on p. 2 of this newsletter.

Also, I would be interested in any arguments any of you want to present for your HOF choices. If you have arguments to advance on behalf of any of the players on the ballot, write them up and send them along with your ballot.

The deadline for voting is as follows: Your ballot must be postmarked no later than February 29, 1992. Vote early--but only once, please.

Hall of Fame Candidates

Dick Allen
Busty Baker
Vida Blue
Bobby Bonds
Ken Boyer
Cesar Cedeno
Orlando Cepeda
John Denny
Rollie Fingers
Curt Flood
Ken Forsch
George Foster
Bobby Grich
Toby Harrah
Jim Kaat
Dave Kingman
Dennis Leonard
Mickey Lolich
Garry Maddox
Bill Mazeroski
Minnie Minoso
Thurman Munson
Ben Oglivie
Tony Oliva
Tony Perez
Vada Pinson
Bill Russell
Ron Santo
Tom Seaver
Rusty Staub
Gorman Thomas
Luis Tiant
Joe Torre
Pete Vukovich
Maury Wills
Steve Yeager