ANALYST

APRIL 1983

ISSUE # 6

IN THIS ISSUE!!!

PAUL SCHWARZENBART LOOKS AT BALLPARKS!
American League this time...............Page 2

QUALITY VERSUS QUANTITY!
Dan Heisman.........................Page 6

RUNS, OR, AND WINS!!
Dallas Adams.......................Page 10

PETE PALMER!!
Adjusting Home Park Factors........Page 13

GUTHBERT MAGNOLIA!
Who? What? .........................Page 17

YOUR SUBSCRIPTION IS UP!!
(If you were a charter subscriber)

$12.00 A YEAR (ONLY)

RE-PRINT
Paul W. Schwarzenbart
Ballpark Effect-Fielding

BALLPARK EFFECTS ON FIELDING STATISTICS—
AMERICAN LEAGUE PARKS

This article summarizes the American League data for the same years as the National League data presented in the earlier article. Baseball Analyst, Issue No. 1, June 1982. Again, the purpose of the study was to examine what effects natural as opposed to astroturf surfaces have on the production of errors and double plays.

As in the earlier National League study, the outstanding conclusion is that astroturf significantly lowers the amount of errors committed.

The years for the study are the same as for the National League study. For 1973, 1976, 1979 and 1980, the statistics include full team/opponent performance at the team home park and on the road. For the 1972 and 1978 seasons only "team" home/road performances were compiled.

The raw statistical tables below show, for example, errors by infielders (E_if), errors by outfielders (E_of) and double plays (DP) in games between Baltimore and their opponents in Baltimore (Games at Home) and at the opponent's home park (Games on Road). This permits a comparison between the percentage of occurrences of double-plays, infield and outfield errors in games between Baltimore and their opponents in games at Baltimore as opposed to games at the opponents' parks.

### TABLE I: RAW STATISTICS

<table>
<thead>
<tr>
<th>Team</th>
<th>Games at Home</th>
<th>Games on Road</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E_if</td>
<td>E_of</td>
</tr>
<tr>
<td>BAL</td>
<td>25.50</td>
<td>6.16</td>
</tr>
<tr>
<td>BOS</td>
<td>26.57</td>
<td>4.72</td>
</tr>
<tr>
<td>CLE</td>
<td>26.27</td>
<td>6.14</td>
</tr>
<tr>
<td>DET</td>
<td>28.39</td>
<td>6.97</td>
</tr>
<tr>
<td>MIL</td>
<td>29.65</td>
<td>6.78</td>
</tr>
<tr>
<td>NY</td>
<td>22.73</td>
<td>6.23</td>
</tr>
<tr>
<td>TOR</td>
<td>24.26</td>
<td>5.11</td>
</tr>
<tr>
<td>CAL</td>
<td>30.72</td>
<td>7.04</td>
</tr>
<tr>
<td>CHI</td>
<td>29.17</td>
<td>7.67</td>
</tr>
<tr>
<td>CHI</td>
<td>27.40</td>
<td>7.44</td>
</tr>
<tr>
<td>KC</td>
<td>27.02</td>
<td>6.30</td>
</tr>
<tr>
<td>MINN</td>
<td>32.12</td>
<td>7.90</td>
</tr>
<tr>
<td>OAK</td>
<td>31.54</td>
<td>7.06</td>
</tr>
<tr>
<td>SEA</td>
<td>26.36</td>
<td>5.18</td>
</tr>
<tr>
<td>TEX</td>
<td>34.45</td>
<td>6.04</td>
</tr>
<tr>
<td>n-KC</td>
<td>28.00</td>
<td>5.25</td>
</tr>
</tbody>
</table>
Before proceeding further, a few remarks on the fields. The Seattle and Toronto data is limited to 2 and one-half seasons, the partial 1978 and full 1979 and 1980 data, for the obvious reason that the teams came into existence in 1977. I set the Kansas City old Municipal stadium data (1972) apart from the other data because I consider the one-half season (Royals home and away) data insufficient for a meaningful analysis. I include it as it shows the general trend but I believe that it is probably unfair to rank Municipal Stadium relative to the others on the basis of such limited data. As for the CHI and CHI-2, if memory serves me correctly, Bill Veeck removed the carpet from Comiskey Park before 1976. I believe that the rug was there from 1971-75. Therefore, the 1972-73 data is for Comiskey as an astroturf park. (If anyone knows the exact dates the Sox played at home on astroturf, please correct me if I'm wrong.)

On to the comparative data. For obvious reasons it is the comparative data which indicates the differences, if any, caused by the parks. If the Yankees, for instance (as Yankee home games show the smallest occurrence of infield errors per 1000 plays, excluding strikeouts), were a good fielding team during this period, there will likely have been fewer errors committed in games involving the Yankees. What we are concerned with is whether more errors were committed in games at Yankee Stadium than on the road. ("Good fielding," in this context refers only to the error-production component of overall fielding ability.)

In the comparative data tables I have simply divided the Games at Home by Games on Road components in order to obtain a ratio of occurrences home/away. The categories are ranked in order of ballpark.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Park</th>
<th>E_{if}</th>
<th>Park</th>
<th>E_{of}</th>
<th>Park</th>
<th>DP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NY</td>
<td>.813</td>
<td>@-SEA</td>
<td>.660</td>
<td>BOS</td>
<td>1.114</td>
</tr>
<tr>
<td>2</td>
<td>@-TOR</td>
<td>.815</td>
<td>@-TOR</td>
<td>.676</td>
<td>OAK</td>
<td>1.086</td>
</tr>
<tr>
<td>3</td>
<td>@-KC</td>
<td>.820</td>
<td>BOS</td>
<td>.691</td>
<td>MIN</td>
<td>1.074</td>
</tr>
<tr>
<td>4</td>
<td>@-CHI</td>
<td>.865</td>
<td>@-KC</td>
<td>.851</td>
<td>TEX</td>
<td>1.069</td>
</tr>
<tr>
<td>5</td>
<td>BOS</td>
<td>.898</td>
<td>CLE</td>
<td>.872</td>
<td>@-CHI</td>
<td>1.053</td>
</tr>
<tr>
<td>6</td>
<td>@-SEA</td>
<td>.928</td>
<td>NY</td>
<td>.909</td>
<td>@-SEA</td>
<td>1.021</td>
</tr>
<tr>
<td>7</td>
<td>CLE</td>
<td>.997</td>
<td>CHI</td>
<td>.916</td>
<td>CLE</td>
<td>1.019</td>
</tr>
<tr>
<td>8</td>
<td>DET</td>
<td>1.030</td>
<td>TEX</td>
<td>.919</td>
<td>CHI</td>
<td>1.007</td>
</tr>
<tr>
<td>9</td>
<td>OAK</td>
<td>1.064</td>
<td>CAL</td>
<td>.960</td>
<td>CAL</td>
<td>.990</td>
</tr>
<tr>
<td>10</td>
<td>CHI</td>
<td>1.070</td>
<td>BAL</td>
<td>1.005</td>
<td>MIL</td>
<td>.989</td>
</tr>
<tr>
<td>11</td>
<td>BAL</td>
<td>1.085</td>
<td>MIL</td>
<td>1.069</td>
<td>DET</td>
<td>.983</td>
</tr>
<tr>
<td>12</td>
<td>MIN</td>
<td>1.086</td>
<td>MIN</td>
<td>1.133</td>
<td>@-KC</td>
<td>.960</td>
</tr>
<tr>
<td>13</td>
<td>MIL</td>
<td>1.093</td>
<td>DET</td>
<td>1.204</td>
<td>BAL</td>
<td>.960</td>
</tr>
<tr>
<td>14</td>
<td>CAL</td>
<td>1.162</td>
<td>OAK</td>
<td>1.224</td>
<td>NY</td>
<td>.953</td>
</tr>
<tr>
<td>15</td>
<td>TEX</td>
<td>1.182</td>
<td>@-CHI</td>
<td>1.635</td>
<td>@-TOR</td>
<td>0.864</td>
</tr>
</tbody>
</table>

|     | KC   | 1.471 | KC   | .658  | KC   | .883|
| Ave.| NAT  | 1.044 | NAT  | .991  | NAT  | 1.022|
|     | AST  | .957  | AST  | .955  | AST  | .975|
First, a few remarks on the tables. The averages referred to are simple numerical averages of the comparative ratios.

Because of the unbalanced numbers of years for certain teams the astroturf and natural turf averages together will not equal one.

As in the National League study, the evidence is overwhelming that astroturf reduces infield errors by a significant margin. The difference in numerical average in NL parks was about 14%; the difference in American League parks is about 19%. If anything the American League data is more overwhelming, the four astroparks are among the top six fielding parks, only Fenway and Yankee stadiums intrude into the astroturf domain. Not surprisingly, the top natural park field in the NL, Los Angeles’ Dodger Stadium is also an exclusively baseball park.

Again, there is again a consistency with the NL results in showing fewer outfield errors on astroturf. In an interesting critique of the NL study, Craig Wright, Texas Rangers Sabermetrician, was surprised at my reluctance to draw a conclusion about the astroturf effect on outfield errors. I surrender! The consistency of the figures is unmistakeable. I was reluctant to draw a conclusion as to outfield errors mainly because of the smaller amount of data, there are far fewer outfield errors committed and it is, therefore, more difficult to establish trends. Wright’s comments and further reflection have persuaded me that the smaller numbers may affect the ability to make comparative conclusions as to individual ballparks but that the overall astroturf/natural turf trend exhibited in the numbers is probably valid. It is, as Wright points out, consistent with common sense.

I also appreciate Wright’s supplemental information concerning double plays based on double play opportunities. Wright mentions two factors which could affect the frequency of double plays per double play opportunities on astroturf and natural turf; (1) prevalence of ground ball on different surfaces; (2) (by implication) speed of the ball’s movement on different surfaces.

When Wright discussed “prevalence” of ground balls on surfaces, I believe he refers to the prevalence of ground ball outs rather than just ground balls, because the slower movement of a ground ball on most natural surfaces means that an infielder of average range will have a better chance of getting to a ball laterally and can play closer to the plate, resulting in shorter throws (most important at third base and shortstop). He suggests that if ground ball outs occur more frequently on natural turf, and that if overall the frequency of double plays on the different surfaces are about equal, as they appear to be, that the double play/opportunity percentage may be higher on astroturf. This suggestion again has some consistency with common sense, although better data would be necessary to discern such a trend. Maybe with the advent of videotapes and home computers such data may soon come available, at least for those teams which televised a lot of games.
Finally, some remarks about the AL data. Relevant factors in the awarding of Gold Gloves were discussed in the NL article. Those factors include fielding percentage, range, hitting, and overall subjective impressions. Certainly another factor is habit. Examining the fielding difficulty of American League parks relative to fielding percentage, the achievements of Brooks Robinson and Mark Belanger are, in retrospect, more impressive.

Based on the number of Gold Gloves awarded in Baltimore over the past 15 or 20 years I had assumed that Baltimore would at least be one the better natural surfaces. At best, it is an average natural surface. Another achievement that stands out is Buddy Bell's recent popularity at the Gold Glove ballot box, playing on a surface that, statistically at least, looks abysmal. I had always thought that Toby Harrah's horrible fielding statistics in Texas had to be in part caused by the playing surface. While Harrah's fielding statistics have improved somewhat in Cleveland, Bell has continued to run up impressive numbers in Texas.

As for future prognostication, players such as Damaso Garcia and/or Griffin and Fernandez (whoever they settle on) in Toronto, the Cruz boys in Seattle, U. L. Washington are astroturf-type infielders playing on astroturf who stand a good chance of winning Gold Gloves. Alan Trammell is a throwback to the Lou Boudreau-type infielder with adequate range on grass, but certainly no Ozzie Smith when it comes to covering ground. Gold Gloves have been rare in the Detroit infield. Robin Yount worked hard to become a better than average shortstop but mainly won the award on the strength of an MVP hitting season, as is often the case when the choice is a close one. Glenn Hoffman and Willie Randolph are good infielders playing on good surfaces who should win a Gold Glove at some time. Playing in New York certainly helped Graig Nettles and Bucky Dent achieve notoriety (deserved or not) for fielding prowess.

In conclusion, ballpark fielding effects should be put in their proper perspective. The outstanding fielder will probably manage to amass, more often than not, impressive fielding statistics even on the worst of fields. It is important, though, to remember that if your favorite shortstop committed 30 or more errors last season, that does not necessarily mean he's a bad fielder, or even that he necessarily has bad hands. That difference of five to ten errors a year that makes an infielder look like a bum may be more due to the conditions than the quality of his play or his ability. Owners and management should likewise be aware of ballpark effects in evaluating the fielding abilities of their players and should be aware of the necessity of selecting players who can thrive on the home playing surface, where, after all, half of the games are played.
Quality versus Quantity

Often asking the correct question is the hardest part of doing a study. One of the questions involving baseball that I like to think about is 'What is 'better' '? This is as in 'Ted Williams was the best hitter of all time' or 'Pete Rose is a better hitter than Richie Ashburn.' What in the world can this mean, if one is trying to be objective?

As a mathematician, I was taught that trying to prove things without axioms to prove them from is futile; what is a proof in one axiom set is garbage in another. Unfortunately, baseball has no such sets of axioms; fans and baseball people make them up and then prove things which others, who have a different set of axioms, will naturally disagree.

One of the axioms in discussions of 'better' is quantity versus quality. This can be a touchy and difficult subject. For instance, it is obvious that in Hall of Fame balloting, quantity is very important. Players such as Koufax and Dizzy Dean are the exceptions rather than the rule. Compare Richie Ashburn's statistics at retirement to Pete Rose's at the same age (both were born near the start of the season, so Rose's stats go through the 1976 season):

<table>
<thead>
<tr>
<th></th>
<th>G</th>
<th>AB</th>
<th>R</th>
<th>H</th>
<th>2B</th>
<th>3B</th>
<th>HR</th>
<th>RBI</th>
<th>BA</th>
<th>SL%</th>
<th>BB</th>
<th>OBP</th>
<th>SB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ashburn</td>
<td>2189</td>
<td>8365</td>
<td>1322</td>
<td>2574</td>
<td>317</td>
<td>109</td>
<td>29</td>
<td>586</td>
<td>.308</td>
<td>.382</td>
<td>1198</td>
<td>.394</td>
<td>234</td>
</tr>
<tr>
<td>Rose</td>
<td>2184</td>
<td>8886</td>
<td>1459</td>
<td>2762</td>
<td>483</td>
<td>101</td>
<td>134</td>
<td>838</td>
<td>.311</td>
<td>.433</td>
<td>957</td>
<td>.378</td>
<td>106</td>
</tr>
</tbody>
</table>

What leaps out at me is that Rose has more power, Ashburn more speed. On items related to how good the team is (At Bats per game, runs and RBI's) Rose far exceeds Ashburn. But in unrelated items, the power versus speed is more noticeable. And, of course, Ashburn gets on base more often due to walks, not unimportant for a leadoff hitter. Yet Ashburn was recently passed by for 15 years by the writer’s committee, while Rose was a first ballot shoe-in long ago. The reasons are many, but one seems to stand out above all the rest: Ashburn didn’t play long enough to establish himself as an 'institution', or one of the 'grand old men' of the game.

So, in trying to answer my question, lets lay down an axiom (although a subjective one at best:) Suppose you were a general manager and you had a crop of rookies to chose from in a pool. IN ADDITION, you had in front of you the ENTIRE CAREER RECORD of each player BEFORE THEY WERE TO PLAY THEIR FIRST MAJOR LEAGUE GAME. Further suppose that you have no other factors, such as team need to use as a selection factor. Then AXIOM #1 is "The better ballplayer is the one a genius general manager would choose in such a situation."

Now this is an intriguing axiom; it pits quantity versus quality in some perspective. For instance, it is clear that one would want to pick Rose over Ashburn, for why not get someone who will produce similarly over 20+ years like Rose rather than only 15. Here, as in Hall of Fame selections, quantity becomes a legitimate comparison factor.

Would you have not have rated Stargell or Phil Niekro or Gaylord Perry lower if they had retired after the 1978 season at ages that would not have been embarrassing? Yet Ashburn never
even saw those ages in his career! Clearly quantity is a large factor in baseball comparisons.

As that general manager, you want to get the most out of each player, and longevity is a factor.

If Kiner hit more home runs per at bat than Hank Aaron, therefore, one could say he was the greater home run hitter; but surely there is value in hitting 755 home runs over 369; many of those extra home runs by Aaron will win ball games that Kiner's replacement couldn't.

There are other problems to consider even with our simple axiom; Rose's team oriented statistics are an example. There is a problem with the axiom; the statistics are inherently biased.

If one leaves in the team oriented values, then one knows how the rest of his 'to be picked' fictitious team will do. For example, let's say one of your rookies is Joe Outfielder, whose lifetime stats are to be a .300 On Base percentage with all singles and walks (no isolated power, no steals) but his final statistics show him averaging 200 runs a year! (This is an example of an 'ad absurdum argument, of course). Should you pick him?? Similarly, but on a lesser scale, Rose's extra At Bats per game, even including Ashburn's walks indicates that your team will be better. So, maybe axiom #1 ought to be modified to not include these factors, such as runs and RBI's and not both games and at bats.

Anyway, now that one has established some axiom, there is much more enjoyment in comparing ballplayers. Of course, the next problem is that players of different eras have different absolute statistics, so maybe we shall limit the information to output per year divided by the 'rague average that year in that statistic. For example, a home run champion in the dead ball era would be given his relative worth, rather than his absolute worth. Now we are getting into an area for which I am going to right another paper about another axiom, which involves comparison of players over eras, but that is another full subject. For now, let us assume the statistics were compiled in the same rough time eras.

Just for fun, let's take pairs of post World War II ballplayers and I'll let you be the general manager. Who would you pick:

Brooks Robinson or Ron Santo; Luis Aparicio or Phil Rizzuto; Carl Yastrzemski or Mickey Mantle
Whitey Ford or Gaylord Perry; Dick Radatz or Ron Reed; Dick Groat or Larry Bowa;
Phil Niekro or Sandy Koufax; Harmon Killebrew or Ralph Kiner; Hank Fidrych or Roger Craig
Don Drysdale or Luis Tiant; Bill Freehan or Roy Campanella; Joe Black or Enrique Romo
Roger Maris or John Mayberry; Duke Snider or Al Kaline; Vernon Law or Doc Medich; Ernie Broglio or Jon Matlack; Catfish Hunter or Don Sutton; Bill Robinson or Wes Covington.

The argument goes on, but at least there are some criteria.

Continuing the question of quality, is it quality per At Bat that is most important, how about per game, per year or per career (sic)? Obviously as one lengthens the time in the denominator, one moves more towards quantity and less towards quality. Just for fun I took some intermediate time (five years) and compiled an interesting set of statistics of a player's best 5 years, not having to be consecutive. I am sure this list tells one something about quality- and maybe even a little about quality. I hope I have added a small amount of knowledge to this subject. - Dan Heisman
<table>
<thead>
<tr>
<th>Name</th>
<th>AB</th>
<th>H</th>
<th>HR</th>
<th>RBI</th>
<th>SB</th>
<th>BA</th>
<th>YR</th>
<th>HR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gehrig</td>
<td>2942</td>
<td>1064</td>
<td>232</td>
<td>850</td>
<td>57</td>
<td>.362</td>
<td>34</td>
<td>30</td>
</tr>
<tr>
<td>Foxx</td>
<td>2870</td>
<td>1000</td>
<td>234</td>
<td>806</td>
<td>30</td>
<td>.348</td>
<td>38</td>
<td>36</td>
</tr>
<tr>
<td>DiMaggio</td>
<td>2726</td>
<td>794</td>
<td>176</td>
<td>706</td>
<td>12</td>
<td>.346</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>Delahanty</td>
<td>2644</td>
<td>1048</td>
<td>56</td>
<td>646</td>
<td>171</td>
<td>.396</td>
<td>93</td>
<td>99</td>
</tr>
<tr>
<td>Cobb</td>
<td>2868</td>
<td>1124</td>
<td>34</td>
<td>550</td>
<td>371</td>
<td>.392</td>
<td>66</td>
<td>65</td>
</tr>
<tr>
<td>Clemente</td>
<td>2829</td>
<td>968</td>
<td>101</td>
<td>465</td>
<td>28</td>
<td>.342</td>
<td>67</td>
<td>70</td>
</tr>
<tr>
<td>Carew</td>
<td>2935</td>
<td>1052</td>
<td>46</td>
<td>387</td>
<td>186</td>
<td>.358</td>
<td>73</td>
<td>77</td>
</tr>
<tr>
<td>Ashburn</td>
<td>2876</td>
<td>975</td>
<td>13</td>
<td>235</td>
<td>117</td>
<td>.339</td>
<td>57</td>
<td>58</td>
</tr>
<tr>
<td>Aaron</td>
<td>2962</td>
<td>975</td>
<td>219</td>
<td>631</td>
<td>56</td>
<td>.329</td>
<td>55</td>
<td>57</td>
</tr>
<tr>
<td>Wagner</td>
<td>2712</td>
<td>979</td>
<td>31</td>
<td>537</td>
<td>243</td>
<td>.361</td>
<td>69</td>
<td>63</td>
</tr>
<tr>
<td>Winzer</td>
<td>2998</td>
<td>1112</td>
<td>42</td>
<td>478</td>
<td>44</td>
<td>.371</td>
<td>69</td>
<td>63</td>
</tr>
<tr>
<td>Williams</td>
<td>2473</td>
<td>916</td>
<td>179</td>
<td>630</td>
<td>10</td>
<td>.370</td>
<td>69</td>
<td>63</td>
</tr>
<tr>
<td>Wilson</td>
<td>2759</td>
<td>914</td>
<td>177</td>
<td>707</td>
<td>33</td>
<td>.331</td>
<td>26</td>
<td>30</td>
</tr>
<tr>
<td>Yastrzemski</td>
<td>2876</td>
<td>877</td>
<td>166</td>
<td>504</td>
<td>67</td>
<td>.305</td>
<td>67</td>
<td>63</td>
</tr>
<tr>
<td>LaJoie</td>
<td>2840</td>
<td>1078</td>
<td>32</td>
<td>523</td>
<td>122</td>
<td>.380</td>
<td>67</td>
<td>77</td>
</tr>
<tr>
<td>Klein</td>
<td>3114</td>
<td>1118</td>
<td>180</td>
<td>693</td>
<td>51</td>
<td>.359</td>
<td>64</td>
<td>65</td>
</tr>
<tr>
<td>Keeler</td>
<td>2808</td>
<td>1107</td>
<td>11</td>
<td>339</td>
<td>231</td>
<td>.394</td>
<td>95</td>
<td>99</td>
</tr>
<tr>
<td>J. Jackson</td>
<td>2833</td>
<td>1076</td>
<td>32</td>
<td>445</td>
<td>135</td>
<td>.380</td>
<td>112</td>
<td>20</td>
</tr>
<tr>
<td>Hornsby</td>
<td>2857</td>
<td>1144</td>
<td>167</td>
<td>664</td>
<td>42</td>
<td>.400</td>
<td>22</td>
<td>29</td>
</tr>
<tr>
<td>Holmman</td>
<td>2774</td>
<td>1071</td>
<td>74</td>
<td>620</td>
<td>40</td>
<td>.336</td>
<td>33</td>
<td>39</td>
</tr>
<tr>
<td>Greenboy</td>
<td>2935</td>
<td>974</td>
<td>201</td>
<td>788</td>
<td>34</td>
<td>.332</td>
<td>35</td>
<td>40</td>
</tr>
<tr>
<td>Speaker</td>
<td>2681</td>
<td>1032</td>
<td>48</td>
<td>505</td>
<td>112</td>
<td>.385</td>
<td>23</td>
<td>20</td>
</tr>
<tr>
<td>Sisler</td>
<td>3062</td>
<td>1154</td>
<td>55</td>
<td>533</td>
<td>166</td>
<td>.377</td>
<td>22</td>
<td>27</td>
</tr>
<tr>
<td>Name</td>
<td>AB</td>
<td>H</td>
<td>HR</td>
<td>RBI</td>
<td>SB</td>
<td>BA</td>
<td>YRS</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>Simmons</td>
<td>2712</td>
<td>1035</td>
<td>131</td>
<td>687</td>
<td>33</td>
<td>.382</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ruth</td>
<td>2594</td>
<td>972</td>
<td>260</td>
<td>766</td>
<td>60</td>
<td>.375</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rose</td>
<td>3226</td>
<td>1081</td>
<td>45</td>
<td>317</td>
<td>49</td>
<td>.335</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F. Robinson</td>
<td>2809</td>
<td>906</td>
<td>203</td>
<td>607</td>
<td>75</td>
<td>.320</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ott</td>
<td>2748</td>
<td>906</td>
<td>173</td>
<td>673</td>
<td>27</td>
<td>.330</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Musial</td>
<td>3042</td>
<td>1090</td>
<td>136</td>
<td>546</td>
<td>30</td>
<td>.358</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mays</td>
<td>2924</td>
<td>954</td>
<td>222</td>
<td>586</td>
<td>122</td>
<td>.324</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mays</td>
<td>2505</td>
<td>823</td>
<td>217</td>
<td>560</td>
<td>62</td>
<td>.329</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In a recent paper\(^*\) this author presented data on the probability of scoring an exact number of runs in a game. One application of this data is to find a formula expressing team Won/Lost percentage as a function of the number of runs the team scores on offense and the number of runs it allows on defense.

The referenced data was compiled from actual major league scores for the 1967 through 1976 seasons, a total of 232 team-seasons, and is reproduced below in Table 1.

To review briefly the data reduction procedure: working season by season, each team was grouped with others which showed a similar runs scored per game average. There were 11 groups in all; the characteristics of each group are included in Table 1. For all the teams in a given group, the number of times each team scored each specific number of runs was tabulated. Once this was completed for all teams, the numbers within each group were converted to percentages. That is, for each group the percentage of times it was shutout was computed, likewise the percentage of games in which it scored exactly one run, etc.

The percentages of Table 1 can be converted to probabilities simply by dividing by 100. For a given group (call it Group A) then, the probabilities of that group making a game score of exactly no runs, exactly one run, etc. are known. Similarly, for any second group, Group B, which is considered to be the opposing team, the probabilities of this opposing team being shutout, held to exactly one run, etc. are also known. Thus it is possible to compute the theoretical winning percentage when any given group plays against any other group.

If \( P_k \) is the probability of Group A scoring exactly \( k \) runs in a game and \( Q_j \) is the probability of Group B scoring exactly \( j \) runs in a game, then the probability of Group A scoring more runs in the game than Group B is:

\[
W = \sum_{k=0}^{18} P_k \sum_{j=0}^{k} Q_j
\]

Similarly, the probability of Group A scoring fewer runs in the game than Group B is:

\[
L = \sum_{j=0}^{18} Q_j \sum_{k=0}^{j} P_k
\]

There remains a certain portion of games in which both groups score the same number of runs. It is assumed that each group goes on to win, in extra innings, the same proportion of those "tie games" as it won of the non-tied games. Thus, ties can be ignored and the won/lost percentage of Group A when playing against Group B is \( W/(W+L) \).

Since the average number of runs scored per game by each group is known from the original data, each group can then be "played" against every other group with the result that a table, see Table 2, of theoretical won/lost percentage (PCT), runs per game (R/G), and opponents' runs per game (OR/G) can be constructed.

By means of least squares regression analysis, some equations giving PCT as a function of R/G and OR/G have been obtained. Three of these

\* The Distribution of Runs Scored
### TEAM WON/LOST PERCENTAGE AS A FUNCTION OF RUNS AND OPPONENTS' RUNS

#### TABLE 1

**GROUP CHARACTERISTICS AND THE PERCENTAGE OF GAMES IN WHICH EXACTLY THE GIVEN NUMBER OF RUNS ARE SCORED**

<table>
<thead>
<tr>
<th>EXACT NUMBER OF RUNS</th>
<th>GROUP 1</th>
<th>GROUP 2</th>
<th>GROUP 3</th>
<th>GROUP 4</th>
<th>GROUP 5</th>
<th>GROUP 6</th>
<th>GROUP 7</th>
<th>GROUP 8</th>
<th>GROUP 9</th>
<th>GROUP 10</th>
<th>GROUP 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>14.52</td>
<td>11.42</td>
<td>9.79</td>
<td>9.98</td>
<td>8.11</td>
<td>7.21</td>
<td>6.49</td>
<td>5.93</td>
<td>4.75</td>
<td>4.01</td>
<td>4.94</td>
</tr>
<tr>
<td>1</td>
<td>18.91</td>
<td>16.90</td>
<td>15.35</td>
<td>13.72</td>
<td>12.90</td>
<td>11.54</td>
<td>10.32</td>
<td>10.00</td>
<td>7.60</td>
<td>7.41</td>
<td>8.02</td>
</tr>
<tr>
<td>6</td>
<td>5.02</td>
<td>6.13</td>
<td>5.72</td>
<td>8.00</td>
<td>7.82</td>
<td>8.40</td>
<td>9.15</td>
<td>9.66</td>
<td>9.64</td>
<td>10.80</td>
<td>9.26</td>
</tr>
<tr>
<td>7</td>
<td>2.72</td>
<td>4.95</td>
<td>4.72</td>
<td>5.12</td>
<td>5.76</td>
<td>6.34</td>
<td>7.00</td>
<td>7.69</td>
<td>7.55</td>
<td>8.95</td>
<td>6.17</td>
</tr>
<tr>
<td>8</td>
<td>1.78</td>
<td>2.15</td>
<td>3.53</td>
<td>3.85</td>
<td>3.72</td>
<td>4.02</td>
<td>5.04</td>
<td>5.49</td>
<td>5.65</td>
<td>5.86</td>
<td>4.94</td>
</tr>
<tr>
<td>9</td>
<td>1.36</td>
<td>1.53</td>
<td>1.68</td>
<td>1.94</td>
<td>2.54</td>
<td>3.14</td>
<td>3.34</td>
<td>3.68</td>
<td>3.52</td>
<td>3.40</td>
<td>7.41</td>
</tr>
<tr>
<td>10</td>
<td>1.25</td>
<td>0.83</td>
<td>0.92</td>
<td>1.27</td>
<td>1.58</td>
<td>2.21</td>
<td>2.21</td>
<td>2.36</td>
<td>3.52</td>
<td>2.78</td>
<td>2.47</td>
</tr>
<tr>
<td>11</td>
<td>0.31</td>
<td>0.53</td>
<td>0.81</td>
<td>0.86</td>
<td>0.95</td>
<td>1.25</td>
<td>1.43</td>
<td>1.81</td>
<td>2.19</td>
<td>3.70</td>
<td>3.70</td>
</tr>
<tr>
<td>12</td>
<td>0.31</td>
<td>0.22</td>
<td>0.33</td>
<td>0.36</td>
<td>0.58</td>
<td>0.68</td>
<td>1.15</td>
<td>1.32</td>
<td>1.24</td>
<td>1.85</td>
<td>1.85</td>
</tr>
<tr>
<td>13</td>
<td>0.21</td>
<td>0.18</td>
<td>0.16</td>
<td>0.24</td>
<td>0.35</td>
<td>0.48</td>
<td>0.53</td>
<td>0.73</td>
<td>0.71</td>
<td>1.54</td>
<td>2.47</td>
</tr>
<tr>
<td>14</td>
<td>0.00</td>
<td>0.00</td>
<td>0.08</td>
<td>0.17</td>
<td>0.25</td>
<td>0.23</td>
<td>0.31</td>
<td>0.54</td>
<td>0.95</td>
<td>0.31</td>
<td>1.24</td>
</tr>
<tr>
<td>15</td>
<td>0.00</td>
<td>0.04</td>
<td>0.08</td>
<td>0.05</td>
<td>0.11</td>
<td>0.20</td>
<td>0.18</td>
<td>0.21</td>
<td>0.43</td>
<td>0.31</td>
<td>0.00</td>
</tr>
<tr>
<td>16</td>
<td>0.10</td>
<td>0.04</td>
<td>0.03</td>
<td>0.02</td>
<td>0.13</td>
<td>0.09</td>
<td>0.05</td>
<td>0.18</td>
<td>0.05</td>
<td>0.62</td>
<td>0.00</td>
</tr>
<tr>
<td>17</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.05</td>
<td>0.11</td>
<td>0.04</td>
<td>0.08</td>
<td>0.10</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>18</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.05</td>
<td>0.01</td>
<td>0.05</td>
<td>0.05</td>
<td>0.08</td>
<td>0.19</td>
<td>0.31</td>
<td>0.00</td>
</tr>
<tr>
<td>19</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.05</td>
<td>0.04</td>
<td>0.05</td>
<td>0.09</td>
<td>0.14</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>20 &amp; more</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.05</td>
<td>0.04</td>
<td>0.05</td>
<td>0.09</td>
<td>0.14</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R/G RANGE</th>
<th>2.75-2.99</th>
<th>3.00-3.24</th>
<th>3.25-3.49</th>
<th>3.50-3.74</th>
<th>3.75-3.99</th>
<th>4.00-4.24</th>
<th>4.25-4.49</th>
<th>4.50-4.74</th>
<th>4.75-5.00</th>
<th>5.00-5.24</th>
<th>5.25-5.49</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO. OF TEAMS</td>
<td>6</td>
<td>14</td>
<td>23</td>
<td>26</td>
<td>49</td>
<td>40</td>
<td>34</td>
<td>24</td>
<td>13</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Equations are listed:

1. \( \text{PCT} = A + B[\log_e(R/G)] \)
   where \( A = 0.034116 \text{OR/G}^2 - 0.3854 \text{OR/G} + 0.8984 \)
   \( B = -0.01809 \text{OR/G}^2 + 0.1497 \text{OR/G} + 0.1233 \)

2. \( \text{PCT} = 0.5 + (R/G - \text{OR/G})[0.1465 - 0.005389(R/G) - 0.006996(\text{OR/G})] \)

3. \( \text{PCT} = 0.5 + 0.4917[\log_e(\text{R/OR})] \)
   where \( \text{R/OR} \) is \( R/G \) divided by \( \text{OR/G} \)

As a check of the accuracy of these equations, eight seasons (one season from each decade since 1901) were selected at random and the formulas were used to calculate the theoretical won/lost percentage for each team. These theoretical percentages were then compared with
actual percentages and standard errors of .025, .026, and .027 were computed for equations 1, 2, and 3, respectively. Thus, equations 1 and 2 are very nearly of equivalent accuracy. Equation 3 is less accurate but compensates by having the advantage of being simpler and faster to use.

**TABLE 2**

**THEORETICAL WON/LOST PERCENTAGES**

<table>
<thead>
<tr>
<th>RUNS PER GAME</th>
<th>OPPONENTS RUNS PER GAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.912</td>
<td>.5000 .4584 .4356 .4099</td>
</tr>
<tr>
<td>3.183</td>
<td>.5416 .5000 .4764 .4494</td>
</tr>
<tr>
<td>3.371</td>
<td>.5644 .5236 .5000 .4719</td>
</tr>
<tr>
<td>3.612</td>
<td>.5901 .5506 .5281 .5000</td>
</tr>
<tr>
<td>3.864</td>
<td>.6190 .5800 .5581 .5289</td>
</tr>
<tr>
<td>4.122</td>
<td>.6422 .6062 .5848 .5557</td>
</tr>
<tr>
<td>4.348</td>
<td>.6666 .6304 .6095 .5806</td>
</tr>
<tr>
<td>4.599</td>
<td>.6869 .6523 .6320 .6037</td>
</tr>
<tr>
<td>4.867</td>
<td>.7116 .6778 .6580 .6292</td>
</tr>
<tr>
<td>5.157</td>
<td>.7322 .7003 .6811 .6534</td>
</tr>
<tr>
<td>5.290</td>
<td>.7431 .7125 .6950 .6666</td>
</tr>
</tbody>
</table>

*The seasons used were 1904 NL, 1916 AL, 1927 AL, 1931 NL, 1948 NL, 1955 AL, 1966 AL, 1976 NL*
ADJUSTED HOME PARK FACTOR

Pete Palmer

Apparent home park factor (HFA) can be measured as follows:

\[
HFA = \frac{\text{Runs scored and allowed}}{\text{Games at home}} - \frac{\text{Runs scored and allowed}}{\text{Games on road}}
\]

However, there are four additional factors which modify this result. First, it is runs per inning that is really desired, not runs per game. Innings can be reliably approximated using available information. Total innings pitched (IP) can be obtained from club pitching records. Total innings batted (IB) can be calculated from club batting statistics. Take total batters faced pitched (at bats plus sacrifices plus walks, hit batters and interferes) less runs scored and runners left on base, all divided by three. As a sidelight, the number of team wins can be found accurately by taking decisions over two plus innings pitched minus innings batted. This is simply because the home team does not bat in the last inning unless the winning run is scored in that inning. This basic fact must be considered in apportioning the innings batted and pitched between home and road. Two more figures are needed, home wins (HW) and road losses (RL).

\[
\begin{align*}
\text{IPH} &= (\text{IP} + \text{RL})/\text{Games-home}/\text{Games-total} \\
\text{IBR} &= (\text{IP} + \text{HW})/\text{Games-road}/\text{Games-total}
\end{align*}
\]

\[
\begin{align*}
\text{IPH} &= \text{IP} - \text{IPH} \\
\text{IBR} &= \text{IP} - \text{IPH}
\end{align*}
\]

A team with home wins equal to road losses will have no correction here, but a typical correction will be on the order of one percent. So the improved home factor (HFI) is calculated as:

\[
HFI = \frac{\text{Runs scored and allowed}}{\text{Inning at home}} - \frac{\text{Runs scored and allowed}}{\text{Inning on road}}
\]

However, true home factor (HFT) includes another correction, touched on by Jim Reuter in the February 1983 Analyst. Since the road park average rating for any team does not include the team's own home park, it is actually slightly to the other side of average from the home factor. This is defined as true road factor (RFT).

Jim's equation, which is correct, relates to the number of teams (NT) in the league.

\[
\frac{\text{HFT}}{\text{NT}} + \frac{\text{RFT(NT-1)}}{\text{NT}} = 1.00
\]

This means that the average home park factors for all parks is equal to one.

What Jim missed was that initial input was HFI, not HFT. By definition:

\[
\text{HFI} = \frac{\text{HFT}}{\text{RFT}}
\]

These last two equations can be solved for HFT as a function of HFI, and then RFT.

\[
\begin{align*}
\text{HFT} &= \frac{\text{NT} \cdot \text{HFI}}{\text{NT} - 1 + \text{HFI}} \\
\text{RFT} &= \frac{\text{HFT}}{\text{HFI}}
\end{align*}
\]

This correction increases as HFI departs from average, and can be over two percent. Since the actual proportion of appearances at home and on the road by individuals is also affected by the fact that the home team doesn't bat in the last inning in games that it won, the individual batting factor (BF) and pitching factor (PF) must be modified.

\[
\begin{align*}
\text{BF} &= \frac{\text{HFT} \cdot \text{IBH/IB} + \text{RFT} \cdot \text{IBR/IB}}{\text{IPH/IP} + \text{RFT} \cdot \text{IPR/IP}}
\end{align*}
\]

This correction is usually less than one percent. To be perfectly accurate, individual day-by-day records could be broken down into home and away categories, rather than assuming that the player breakdown was the same as the team.
A final correction, which is not related to home park but is related to individual performance is that batters on a particular team do not have to face pitchers on that team. To begin this calculation, team bat rating (TBR) and team pitch rating (TPR) are found, neglecting the fact that they don't have to face each other.

\[
TBR = \frac{\text{Runs scored-team}}{\text{BF} \cdot \text{Runs-league/innings-league}}
\]

\[
TPR = \frac{\text{Runs allowed-team}}{\text{PF} \cdot \text{Runs-league/innings-league}}
\]

Next, this fact is included. Since each rating depends on the other, these equations can be repeated, using as input the output of the previous iteration. However, two or three iterations should be sufficient.

\[
TBR' = TBR \cdot \frac{(NT - 1)}{(NT - TPR')}
\]

\[
TPR' = TPR \cdot \frac{(NT - 1)}{(NT - TBR')}
\]

This factor also becomes larger for team with ratings removed from average, and can be in the one to two percent range. These factors can be combined with the batting and pitching factors to give and overall factor.

\[
OBF = BF \cdot \frac{(NT - TPR')}{(NT - 1)}
\]

\[
CPF = PF \cdot \frac{(NT - TBR')}{(NT - 1)}
\]

In my original research, I calculated home park factors by using runs scored and allowed. However, I have since decided that it is fairer to use just runs allowed. This is because the home team may be made up of batters who can take advantage of the particular park and thus make it appear to be a better hitting park than it should be. This could also be true of pitchers, but I don't believe that pitchers are anywhere near as apt to be chosen in that manner. The equations in this article still apply, except that in the HFA and HFT equations, these factors are divided by average home/road figures. Looking at the results for all teams from 1977 through 1982, there are only two cases in which the home park factor using only runs allowed differs from the overall park factor by more than four percent. Kansas City shows .97 with pitchers only, 1.02 counting both, while Houston shows .80 with runs allowed, .86 with both. In 1981, Houston allowed only 106 runs in 51 home games, the lowest in major league history. Strangely enough, the best hitting park in each league shows about the same park factor in either way, indicating that neither team has succeeded in tailoring its lineup. Fenway Park was 1.17 either way, while Atlanta was 1.16 for the opponents, 1.17 overall.

The standard deviation for a six-year park factor can be found using a reasonable approximation for baseball scoring. If runs were scored singly like goals in a hockey game, the standard deviation would be equal to the square root of the number of runs involved. However, since in baseball, runs can come in groups, it turns out that every actual distribution of runs I have seen has a standard deviation equal to the square root of twice the number of runs involved. These six-year averages involve approximately 4000 runs at home and 4000 on the road, so the standard deviation is \(\sqrt{2 \cdot 4000}\) or 126 over 4000 or about 3 percent. Using just runs allowed, you get \(\sqrt{2 \cdot 4000}\) or 89 over 2000 or about 4.5 percent. Thus the error is increased by using a smaller sample, but the mean value should be more accurate. Even the large differences at Kansas City and Houston are not unexpected considered the size of the standard deviation. The fact that runs have a standard deviation equal to twice the square root of the number of runs can be substantiated using page two from Adams' Distribution of Runs Scored-in issue number one of the Analyst.
For an example, take Atlanta in 1977. The figures were:

at home - 81 games, 40 wins, 416 runs scored, 488 runs allowed
on road - 81 games, 60 losses, 262 runs scored, 407 runs allowed
batters faced pitcher - (5534 ab + 83 sh + 34 sf + 537 bb + 17 hb + 1 int)
innings batted - (6206 bfp - 678 r - 1139 lob )/3 = 1463
innings pitched - 1445

\[
HFA = \frac{904}{81} = 1.351
\]
This would lead to an apparent bat factor
\[
\frac{669}{81} = (1.135 + 1)/2 = 1.18
\]

\[
IPH = \frac{(1445 + 60)/2}{2} = 752
\]

\[
IBR = \frac{(1463 + 40)/2}{2} = 751
\]

\[
HFI = \frac{904/(752 + 712)}{669/(693 + 751)} = 1.333
\]

\[
HFT = 12 \cdot 1.333 = 1.297
\]

\[
RFT = \frac{1.297}{1.333} = .973
\]

\[
BF = 1.297 \cdot \frac{712/1463 + .753 \cdot 751/1463 = 1.131}{11 + 1.333}
\]

\[
PF = 1.297 \cdot \frac{752/1445 + .753 \cdot 693/1445 = 1.142}{11 + 1.333}
\]

\[
TBR = \frac{678/1463}{1.131 \cdot .488} = .840
\]

\[
TPR = \frac{895/1445}{1.142 \cdot .488} = 1.111
\]

\[
TBR' = .840 \cdot \frac{11/(12-1.111)}{1.111} = .849
\]

\[
TPR' = 1.111 \cdot \frac{11/(12-1.095)}{1.111} = 1.095
\]

\[
TBR'' = .840 \cdot \frac{11/(12-1.095)}{1.111} = .847
\]

\[
TPR'' = 1.111 \cdot \frac{11/(12-1.095)}{1.111} = 1.096
\]

\[
OBF = 1.131 \cdot \frac{(12-1.096)/11}{1.121}
\]

\[
OFF = 1.142 \cdot \frac{(12-1.095)/11}{1.158}
\]

Thus the overall bat factor is 1.12, quite a bit lower than the apparent 1.18

In Jim's example, an HFI of 1.20 would lead to an HFT of 1.180, RFT of .983
and a BF of 1.082, assuming innings batted at home and away were the same.
Assuming a typical home win percentage of .535, the figure is reduced to 1.079.

Using just runs allowed, the overall league data is needed.
4135 runs allowed at home, 4421 runs allowed on the road
9031 innings pitched at home (17515 innings + 547 road losses)/2
8484 innings pitched on the road (17515-9031)

\[
\text{league average runs allowed per inning at home} = \frac{4135}{9031} = .879
\]

\[
\text{runs allowed per inning on road} = \frac{4421}{8484}
\]

\[
HFA = \frac{488}{752} = 1.257
\]

\[
HFT = 1.231
\]

\[
TBR = .862
\]

\[
OBF = 1.087
\]

Thus the fact that the Braves' batters were better able to take advantage
of their home park than most teams does not penalize them. The final park
and staff adjustment is only half the apparent park factor.
Here is data for all major league parks covering the 1977-82 period.

<table>
<thead>
<tr>
<th>Team</th>
<th>HOME RUNS</th>
<th>RUNS ALLOWED ONLY</th>
<th>RUNS SCORED+ALLOWED</th>
<th>ALLOWS</th>
<th>HIT+ALLOWS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ALLOWS</td>
<td>HFT TBR TPR OBF OPF</td>
<td>HFT TBR TPR OBF OPF</td>
<td>HFT TBR TPR OBF OPF</td>
<td>HFT TBR TPR OBF OPF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HFT TBR TPR OBF OPF</td>
<td>HFT TBR TPR OBF OPF</td>
<td>HFT TBR TPR OBF OPF</td>
<td>HFT TBR TPR OBF OPF</td>
</tr>
<tr>
<td>Balt</td>
<td>.88 .109  .96 .95 .94</td>
<td>.92 .106  .94 .97 .96</td>
<td>1.06 .116  .94 .98 .120 .98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bos</td>
<td>1.17 1.03  .93 1.08 1.08</td>
<td>1.17 1.03  .93 1.08 1.08</td>
<td>1.15 1.18 1.00 1.19 1.16 .98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cal</td>
<td>.96 1.06 1.02 .98 .98</td>
<td>.95 1.06 1.02 .98 .97</td>
<td>.97 1.05 .98 .97 .105 .97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi</td>
<td>1.93 1.00 1.01 1.01</td>
<td>1.99 1.01 1.02 .99 .99</td>
<td>.82 1.04 .94 .82 1.04 .95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clev</td>
<td>.97 1.00 1.07 .98 .98</td>
<td>.99 1.00 .96 .99 .100</td>
<td>.93 .79 .97 .103 .75 .90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Det</td>
<td>1.06 1.00 1.06 1.02 1.02</td>
<td>1.06 1.00 1.06 .103 1.03</td>
<td>1.27 .98 1.01 1.28 .97 1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KC</td>
<td>.97 1.10 1.00 .99 .93 1.02 1.08 .97 1.01 1.00</td>
<td>1.07 .94 1.01 1.01 .97 1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milw</td>
<td>.98 1.10 1.00 .99 .98</td>
<td>.95 1.12 1.01 .98 .97</td>
<td>.84 1.42 1.12 .84 1.42 1.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minn</td>
<td>1.03 1.04 1.01 1.02 .92</td>
<td>1.06 .96 1.00 1.03 1.03</td>
<td>1.00 .80 1.01 1.05 .78 .99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NY</td>
<td>.94 1.07 1.02 1.02 .98</td>
<td>.94 1.07 1.02 .98 .97</td>
<td>.91 1.23 .90 .94 1.21 .89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oak</td>
<td>.89 1.07 1.07 .94 .95</td>
<td>.90 1.02 1.07 .95 .96</td>
<td>1.07 1.00 1.11 .89 .99 1.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sea</td>
<td>1.05 1.08 1.08 1.02 1.02</td>
<td>1.05 .88 1.08 1.02 1.03</td>
<td>1.38 .75 1.00 1.44 .74 .98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tex</td>
<td>.96 1.00 1.00 1.00 1.00</td>
<td>1.01 1.01 1.00 1.00 1.00</td>
<td>.95 1.93 .92 .89 .96 .95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tor</td>
<td>1.11 1.02 1.02 1.05 1.07</td>
<td>1.08 1.01 1.03 1.03 1.05</td>
<td>1.08 .74 1.02 1.05 .75 1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl</td>
<td>1.16 1.02 1.03 1.07 1.08</td>
<td>1.17 1.02 1.03 1.07 1.09</td>
<td>1.69 .89 1.00 1.69 .89 .90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi</td>
<td>1.14 1.02 1.01 1.06 1.07</td>
<td>1.18 1.01 1.00 1.08 1.10</td>
<td>1.29 .81 1.02 1.30 .81 .92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cin</td>
<td>1.04 1.03 1.00 1.02 1.02</td>
<td>1.02 1.04 1.01 1.01 1.00</td>
<td>1.12 1.07 1.01 1.07 1.09 1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hou</td>
<td>.80 1.02 1.02 1.02 1.02</td>
<td>1.00 1.04 1.01 1.01 1.00</td>
<td>1.07 1.07 1.07 1.07 1.07 1.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LA</td>
<td>.94 1.09 1.09 1.09 1.09</td>
<td>.94 1.09 1.09 1.09 1.09</td>
<td>1.14 1.31 1.31 1.31 1.31 1.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mon</td>
<td>.99 1.02 1.02 1.02 1.02</td>
<td>.99 1.02 1.02 1.02 1.02</td>
<td>1.05 1.25 1.25 1.25 1.25 1.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NY</td>
<td>.99 1.03 1.03 1.03 1.03</td>
<td>.99 1.03 1.03 1.03 1.03</td>
<td>1.04 .72 1.04 1.01 .73 1.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phi</td>
<td>1.00 1.09 1.00 1.00 1.00</td>
<td>1.03 1.07 1.08 1.02 1.01</td>
<td>1.09 1.11 1.12 1.12 1.12 1.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pit</td>
<td>1.09 1.01 1.01 1.01 1.01</td>
<td>1.03 1.02 1.03 1.04 1.04</td>
<td>1.16 1.04 1.12 1.12 1.12 1.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>StL</td>
<td>1.04 1.03 1.02 1.02 1.02</td>
<td>1.01 1.04 1.01 1.00 1.00</td>
<td>.80 .86 1.03 .97 .87 1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.86 1.01 1.09 1.09 1.09</td>
<td>.85 1.01 1.10 .93 .93</td>
<td>.75 .83 1.12 .74 .84 1.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF</td>
<td>.95 1.08 1.02 1.02 1.02</td>
<td>.93 .99 1.03 .97 .97</td>
<td>.78 1.14 1.06 .76 1.16 1.08</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A number greater than one in the TBR column indicates good batting, one less than one in the TBR column indicates good pitching. It is interesting to note that with the park adjustment, the Red Sox have the second best pitching in the majors over the past six years, while their batting ranks only ninth. Houston's batting is slightly better than average, and almost as good as Boston's, while Atlanta and Chicago NL rank close to the bottom of the league in batting. The home run data was calculated in the same manner.

My data file includes games, wins, losses, ties, runs scored and allowed and homers hit and allowed at home, away and total for every major league team from 1900 to the present. In addition, I have individual data on almost every prominent American League player, about 150 in all. The individual data and American League park data was reported in the 1978 Baseball Research Journal.
Dear James Bill,

I been readin this here Baseball Analyst for almost a year now and I aint ever read such pure twaddle, sheer foolishness, absolute falderal and utter garbage. Its time all these rank amateur Savymetricians cleared the floor and let a real pro take over.

I do kinda take to your method of Runs Produced. Trouble is, you got no similar method for the pitchers. I'll pull your fat out of the fire, boy. Just you harken up to the next few paragrafs and I'll smarten up your readers --- if you got any left.

Lets start with the basics. All pitchers is either starters or relievers. Now, I grant you some few (12-15%?) is both during any given season. Usually these starter-reliever-part-timers are young unestablished pitchers so they wont inter­fer none with what Im proposin. All of us knowledjubul students of the game know that the most important measurement of a starting pitcher is how many men he puts on base (BRA). Walk to strikeout ratios are all well and good but a 2:1 SO to W ratio dont guarantee success --- Duggy Bird, Frank Di Pino, Fowlkes, Hammaker, Tugo McGraw, Donnie Moore, Bert Roberge, Ruhle, Tom Brennan, Moose Haas, Jon Matlack, Gaylord Old Man, Shane Rally, El Tante all had poor to indifferent 1982s while lookin slick in this area. Hits to innings pitched falls into the same misery --- lots look good here but walk so many failers that they get zonked when they can least afford to.

Relief pitchers on the other hand often stroll on the scene when the cards is stacked agin em. If they gotta pitch around somebody they dont care if they gotta walk em. BRA dont mean jack shit in their case. The one thing thats gonna burn their derriere is base hits so this is the best measurement for these good ol boys. All you gotta do is check out the stats of Dom Stanhouse and Butch Metzger when these guys was in Hog Heaven to get my point.

By figurin the BRA for all AL starters we come up with a figure of 1.391 --- any pitcher who has more than 1/3 of his game appearances as starts is put in this category. The BRA for the NL is 1.341. The H:IP ratio is .972 for AL relievers and .934 for NL boys.

Everbody follow me so far? So good. Once a pitcher ratio is added or subtracted to or from the league ratio were hot to trot. Then we multiply this figure by his percentage of his teams total inning pitched in his respectable category

Example follows on next page for those who aint got too many smarts.
What really grabs you about this system is that it determines the quality element by comparing the pitchers performance with the league average and then extends its value by comparing the quality factor with members of his own team to show the participation factor. BRA performance alone won't cut it. If a guy works in 45 innings and has a BRA of plus .412 it sure don't measure up to the other guy who works in 250 innings and gas a BRA of .400. So % of a feller's team appearances is more important than the number of games in which he shows up. Why? Well, number of games appearing in is great for a shorty reliever but plays hob with the long boys. After all, it's easier to pitch 1 inning in 3 of 5 days than it is to pitch 3 innings in 3 of five days and yet under the game appearance criterion the shorty gets more credit.

Well, that's it. Just check out the Top 5 for 1982 below. Then I'll show you how the teams stack up.

**National League**

<table>
<thead>
<tr>
<th>Starters</th>
<th>IP</th>
<th>BRA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joe Niekro</td>
<td>64</td>
<td>58</td>
</tr>
<tr>
<td>Joaquin Andujar</td>
<td>61</td>
<td>56</td>
</tr>
<tr>
<td>Mario Soto</td>
<td>60</td>
<td>55</td>
</tr>
<tr>
<td>Steve Wonder</td>
<td>55</td>
<td>54</td>
</tr>
<tr>
<td>Reuss</td>
<td>51</td>
<td>53</td>
</tr>
<tr>
<td>Rogers</td>
<td>49</td>
<td>52</td>
</tr>
<tr>
<td>Tortilla Fats</td>
<td>44</td>
<td>51</td>
</tr>
<tr>
<td>Sutton</td>
<td>42</td>
<td>50</td>
</tr>
<tr>
<td>Lea</td>
<td>33</td>
<td>49</td>
</tr>
<tr>
<td>Christensen</td>
<td>30</td>
<td>48</td>
</tr>
</tbody>
</table>

**American League**

<table>
<thead>
<tr>
<th>Starters</th>
<th>IP</th>
<th>BRA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dave Stieb</td>
<td>68</td>
<td>58</td>
</tr>
<tr>
<td>Jim Palmer</td>
<td>66</td>
<td>57</td>
</tr>
<tr>
<td>Clancy</td>
<td>60</td>
<td>56</td>
</tr>
<tr>
<td>Barker</td>
<td>40</td>
<td>55</td>
</tr>
<tr>
<td>Eckersley</td>
<td>40</td>
<td>54</td>
</tr>
<tr>
<td>Bannister</td>
<td>38</td>
<td>53</td>
</tr>
<tr>
<td>Hoyt</td>
<td>32</td>
<td>52</td>
</tr>
<tr>
<td>John</td>
<td>31</td>
<td>51</td>
</tr>
<tr>
<td>Sutcliffe</td>
<td>30</td>
<td>50</td>
</tr>
<tr>
<td>Caldwell</td>
<td>30</td>
<td>49</td>
</tr>
</tbody>
</table>

Pete Vuckovich - the Cy Young winner - finished up with a sub-mediocre - 23. Selecting him was like naming my 85-year-old maiden aunt, Tilly Grossbeak, Miss County Seat of Turnup Green, Mississippi. Which actually happened. Everybody knows either Mr. Jockey Shorts or Stieb shoulda been the winner.

Before we get to the teams let me silence all you well-meaning but untutored critics. Your gonna say "Your system don't allow for those guys who pitch in both relief and the begimmun. So what. How many of those dudes are there? And if you care that much go back to the box scores and figure it out. Do you want me to do all the work?
THE CLUBS

NATIONAL LEAGUE

<table>
<thead>
<tr>
<th>Starters</th>
<th>Relievers</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlanta</td>
<td>- 36</td>
<td>25</td>
</tr>
<tr>
<td>Chicago</td>
<td>- 60</td>
<td>- 16</td>
</tr>
<tr>
<td>Cincinnati</td>
<td>- 37</td>
<td>46</td>
</tr>
<tr>
<td>Houston</td>
<td>109</td>
<td>9</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>118</td>
<td>13</td>
</tr>
<tr>
<td>Montreal</td>
<td>98</td>
<td>- 63</td>
</tr>
<tr>
<td>The Nutts</td>
<td>- 138</td>
<td>- 28</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>- 25</td>
<td>- 75</td>
</tr>
<tr>
<td>Pittsburgh</td>
<td>- 33</td>
<td>28</td>
</tr>
<tr>
<td>St. Louis</td>
<td>16</td>
<td>40</td>
</tr>
<tr>
<td>San Diego</td>
<td>42</td>
<td>73</td>
</tr>
<tr>
<td>San Francisco</td>
<td>- 98</td>
<td>- 9</td>
</tr>
</tbody>
</table>

AMERICAN LEAGUE

<table>
<thead>
<tr>
<th>Starters</th>
<th>Relievers</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baltimore</td>
<td>79</td>
<td>25</td>
</tr>
<tr>
<td>Boston</td>
<td>- 92</td>
<td>18</td>
</tr>
<tr>
<td>California</td>
<td>75</td>
<td>38</td>
</tr>
<tr>
<td>Chicago</td>
<td>- 12</td>
<td>51</td>
</tr>
<tr>
<td>Cleveland</td>
<td>- 16</td>
<td>12</td>
</tr>
<tr>
<td>Detroit</td>
<td>48</td>
<td>15</td>
</tr>
<tr>
<td>Kansas City</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>Milwaukee</td>
<td>- 18</td>
<td>3</td>
</tr>
<tr>
<td>Minnesota</td>
<td>- 67</td>
<td>2</td>
</tr>
<tr>
<td>New York</td>
<td>6</td>
<td>58</td>
</tr>
<tr>
<td>Oakland</td>
<td>- 79</td>
<td>- 94</td>
</tr>
<tr>
<td>Seattle</td>
<td>19</td>
<td>48</td>
</tr>
<tr>
<td>Texas</td>
<td>- 36</td>
<td>- 114</td>
</tr>
<tr>
<td>Toronto</td>
<td>58</td>
<td>- 3</td>
</tr>
</tbody>
</table>

Well, there it is. Prove positive of the results. Now, all you smarty-pants Savvy-metricians can go back to the drawing boards and beat your brains out to come up with somethin better. I challenge you.

Scuse-me, I forgot the relievers.

<table>
<thead>
<tr>
<th>National League</th>
<th>American League</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bedrosian</td>
<td>The Goose</td>
</tr>
<tr>
<td>DeLeon</td>
<td>Caudill</td>
</tr>
<tr>
<td>Bair</td>
<td>Armstrong</td>
</tr>
<tr>
<td>Reardon</td>
<td>R L Jackson</td>
</tr>
<tr>
<td>Scurry</td>
<td>Van de Berg</td>
</tr>
<tr>
<td>Orosco</td>
<td>T Martinez</td>
</tr>
<tr>
<td>Forster</td>
<td>Felton</td>
</tr>
<tr>
<td>Kern</td>
<td>Fingers</td>
</tr>
<tr>
<td>Garber</td>
<td>J McLaughlin</td>
</tr>
<tr>
<td>Lesley</td>
<td>Andy Hassle</td>
</tr>
</tbody>
</table>

Dont this all look purty? Eat you hearts out - you great logishuns.

Your pal,

Cuthbert Magnolia
Just a couple of quick points from the editor:

1) I strongly disagree with Dan Heisman's comment that "It is obvious that in Hall of Fame balloting, quantity is very important. Player's such as Koufax and Dizzy Dean are the exception rather than the rule." On the contrary, it is only in the last fifteen years that career stats have come to mean anything at all in Hall of Fame selections, and to this day it is seasonal, not career, numbers that define Hall of Fame membership.

The Hall of Fame is rife with players, like Koufax, Dean, Jackie Robinson, Hack Wilson, Klein, Campanella, Bill Terry, Lindstrom, Waddell, Chesbro, Chance, Baker, Joss, and many others who were outstanding players for maybe five years, and did not compile impressive career totals. Many players had a longer, flatter career spiral and compiled bulkier totals, like Bunning, Pinson, Doc Cramer, Aparicio, Fox, Wilhelm, and Quinn. And in other cases, players like Goose Goslin, Roger Conner, Sam Crawford and Zack Wheat had to wait for Hall of Fame selection until after contemporaries who had shorter, more meteoric careers.

In few cases is there really a conflict, because in every generation those players who reach the highest peaks and those who have the longest careers are pretty much the same players. The correlation between peak value and career length is extremely high. But when there is a conflict, it is peak value, not gross value, which is the dominant criterion.

2) I choose a different option than Pete Palmer, I believe, in defining what it is we are looking for when we adjust for ballpark effects. My feeling is that I am adjusting not for the "true ability" of the player, which is a thing that we can never know, but for the value of the player in the most relevant circumstances.

For that reason, I find Pete's individual park adjustments, while certainly interesting, not germane to the question of the player's value. Suppose that you have a player who plays in a pitcher's park, but who, because of his ability to take advantage of some unique feature of the park, hits extremely well there. We might be talking about a player who is a dead-pull power hitter playing in Baltimore. Should that player's run estimates be reduced when adjusting him to a normal park? Absolutely not. Wins are the goal of the team, and therefore wins define real value for the team. If the player creates 100 runs while playing in Baltimore, those 100 will likely result in more wins for his team than 100 runs in Fenway.

Thus I disagree with Pete's decision that "it is fairer to (base the adjustments) on just runs allowed, ...because the home team may be made up of batters who can take advantage of the particular park and this might make it appear to be a better hitting park than it should be." No, I say. I say, if they can take advantage of the park, then it is a hitter's park--indeed, that is exactly what a hitter's park is, a park which allows hitters to take advantage of it. What the park "should be" is something we can never know.

The process of adapting talents to parks is a two-sided one, offensive and defensive. I see no advantage whatever in ignoring one of those.

Consider an economic analogy. Suppose inflation raises the cost of goods by 10%, and raises average salaries by 10%. Your own salary, however, does not go up. Should we say, in adjusting for your personal economic situation, that you have not been affected by the inflation? Of course not. But that is exactly what has happened. If Fenway increases runs by 10%, it raises the cost of a win by 10%. When we adjust for the "worth" or "value" of the player, we must adjust not for the impact of the park upon any individual, nor for the impact of the park upon one side of the equation. We must adjust for its impact upon the entire on-field economy.