
By the Numbers

Volume 21, Number 1

The Newsletter of the SABR Statistical Analysis Committee

November, 2011

Comment

Academic Research: JQAS 6.2

Charlie Pavitt

The author reviews two recent baseball papers from the Journal of Quantitative Analysis in Sports.

Rosenfeld, Jason W., Jake I. Fisher, Daniel Adler, and Carl Morris, Predicting overtime with the Pythagorean formula, Journal of Quantitative Analysis in Sports, Vol. 6 Issue 2, Article 1

The Pythagorean formula is conventionally applied to computing a team's winning percentage based on the number of runs scored and given up by that team. In an imaginative move, Rosenfeld, Fisher, Adler, and Morris (2010) extended its application to predicting the odds that a team will win an extra-inning game.

The authors used data from 2000 through 2008, during which 8.1 percent of the games (1775) went into an average of 2.1 extra innings. The home team won about 53 percent of the time, a tick lower than the usual 54, and the author(s) adjusted their estimation for home field advantage based on that finding (has anyone used Pythagoras separately for home and away games?). They made the obvious-when-you-think-about-it point that the exponent will be smaller than 2 because the shorter length of almost all extra inning games, in their words, creates more randomness; another way to put it is that the inherent talent difference between teams is less likely to assert itself the fewer innings the teams play.

Indicative of that point, they uncovered an exponent of 1.94 for

normal-length games but one of only .94 for extra innings. The ratio between the two (.48) is almost exactly what Rosenfeld et al expected given that the square root of 2.1 divided by 9, in essence an estimation of relative randomness, is .5. According to their estimates, a .700 team would win about 60 percent of their extra inning games and a .750 team (if there ever were one) about 63 percent.

In this issue

Academic Research: JQAS 6.2.....	Charlie Pavitt.....	1
Was the Power Turned Off in the 1980s?.....	Donald A. Coffin.....	3

Note

This issue of "By the Numbers" immediately follows the February, 2010 issue (20.1). We are jumping forward in order to bring the issue date closer to the publication date.

Addona, Vittorio and Jeremy Roth, Quantifying the effect of performance-enhancing drug use on fastball velocity in major league baseball, Journal of Quantitative

Analysis in Sports, Vol. 6 Issue 2, Article 6

The subject matter of this piece is obvious from its title. The author's dataset included average fastball velocity (as measured by Baseball Info) for the years 2002 through 2008. Using a cross-sectional design ignoring attrition rates, they noted an average fastball speed of 90.10 miles per hour, with a gradual decrease from 92½ for 20 year olds to about 87½ for 40 year olds.

As they admit, there is clearly a selection bias here, as harder-throwing pitchers tend to reach the majors at an earlier age; they cite other evidence suggests an inverted-U relationship with a peak at about 26 or 27. Interestingly, they also observed an

asymptotically increase in fastball velocity across months, jumping from 89.79 in April to 90.07 in May to 90.25 in June, with a slower rise up to 90.37 in September. Turning to their main purpose, Addona and Roth contrasted from the main data set those pitchers who were either suspended for PED use starting with 2005 or were mentioned in the Mitchell report; in total, 27 pitchers accused of steroid use and 16 of human growth hormone.

The results of three different methods for guesstimating when pitchers were and were not using resulted in the following conclusions: HGH has either no or a slightly negative impact on fastball pitch speed, but steroids have a positive effect estimated as averaging about one mile per hour that widens as pitchers age, particularly for relievers. This conclusion is consistent with those for hitters; steroids increases performance but HGH has no impact.

The papers reviewed here can be downloaded at <http://www.bepress.com/jqas/vol6/iss2/>.

Charlie Pavitt, chazzq@UDel.Edu ♦

“By the Numbers” mailing list

SABR members who have joined the Statistical Analysis Committee will receive e-mail notification of new issues of BTN, as well as other news concerning this publication.

The easiest way to join the committee is to visit <http://members.sabr.org>, click on “my SABR,” then “committees and regionals,” then “add new” committee. Add the Statistical Analysis Committee, and you’re done. You will be informed when new issues are available for downloading from the internet.

If you would like more information, send an e-mail to Neal Traven, at beisbol@alumni.pitt.edu. If you don’t have internet access, we will send you BTN by mail; write to Neal at 4317 Dayton Ave. N. #201, Seattle, WA, 98103-7154.

Submissions

Phil Birnbaum, Editor

Submissions to *By the Numbers* are, of course, encouraged. Articles should be concise (though not necessarily short), and pertain to statistical analysis of baseball. Letters to the Editor, original research, opinions, summaries of existing research, criticism, and reviews of other work are all welcome.

Articles should be submitted in electronic form, preferably by e-mail. I can read most word processor formats. If you send charts, please send them in word processor form rather than in spreadsheet. Unless you specify otherwise, I may send your work to others for comment (i.e., informal peer review).

I usually edit for spelling and grammar. If you can (and I understand it isn’t always possible), try to format your article roughly the same way BTN does.

I will acknowledge all articles upon receipt, and will try, within a reasonable time, to let you know if your submission is accepted.

Send submissions to Phil Birnbaum, at birnbaum@sympatico.ca.

Was the Power Turned Off in the 1980s?

Donald A. Coffin

A recent article in "Baseball Research Journal" looked at league-leading home run numbers to argue that power was significantly down in the decade of the 1980s. Here, the author responds that if you look at all players, not just those at the top of the rankings, the 1980s weren't as light-hitting as they may otherwise appear.

Introduction

In the most recent issue of the *Baseball Research Journal*,¹ Dan D'Addona argues that the 1980s were a decade in which "...power numbers were noticeably down, making the eye-opening numbers of the steroid-enhanced decades following look even gaudier" (p. 66). His evidence is the relatively small number of players-seasons in which 40 or more home runs were hit (13) in the 1980s, which was "...the lowest of any decade since Babe Ruth revolutionized the game" (p. 66).

Table 1 shows clearly that, in that respect, the National League-leading totals were, in the 1980s, considerably lower than in earlier or later decades. In the NL, the three earlier decades (1950s, 1960s, and 1970s) were quite similar, with league-leading totals averaging around 44 HR per year. The two succeeding decades were also similar, with league-leading totals averaging about 48 in the 1990s and about 51 in the 2000s. The pattern in the American League, however, is very different. The 1960s (league-leading totals averaging about 46), the 1990s, and the 2000s (about 48 in both decades) are quite comparable. The 1950s (about about 39) and the 1970s (about 37) are comparable. In the AL, the 1980s are between the low-power 1950s and 1970s and the high-power 1960s, 1990s, and 2000s (with league-leading totals averaging about 42 in the 1980s).²

Table 1 – League-Leading Home-Run Totals (Average by Decade)

Decade	NL	AL
1951–1960	44.7	39.5
1961–1970	44.7	46.5
1971–1980	43.2	37.3
1982–1990	40.2	42.1
1991–2000	48.3	48.2
2001–2010	51.3	48.5

But power is not just something that manifests itself in league-leading totals (granted that those are most widely noticed), but is something that may appear throughout the hitting performances of all players in the major leagues. And power is not just home runs, but can also show up in doubles, and in Isolated Power (Slugging Average minus Batting Average). So we can ask whether any downturn in home-run hitting reflects a more general decline in power statistics or not, by looking at league-leading doubles totals, at doubles per plate appearance, at home runs per plate appearance, and at Isolated Power.

¹ Dan D'Addona, "Baseballs Forgotten Era: Decade Saw Lowest Individual Power Since Home Runs Emerged With Ruth," *Baseball Research Journal*, V. 40, N. 2, pp. 66-68.

² All data are from Baseball Reference (http://www.baseball-reference.com/leaders/HR_leagues.shtml).

Doubles Power: League-Leading Totals³

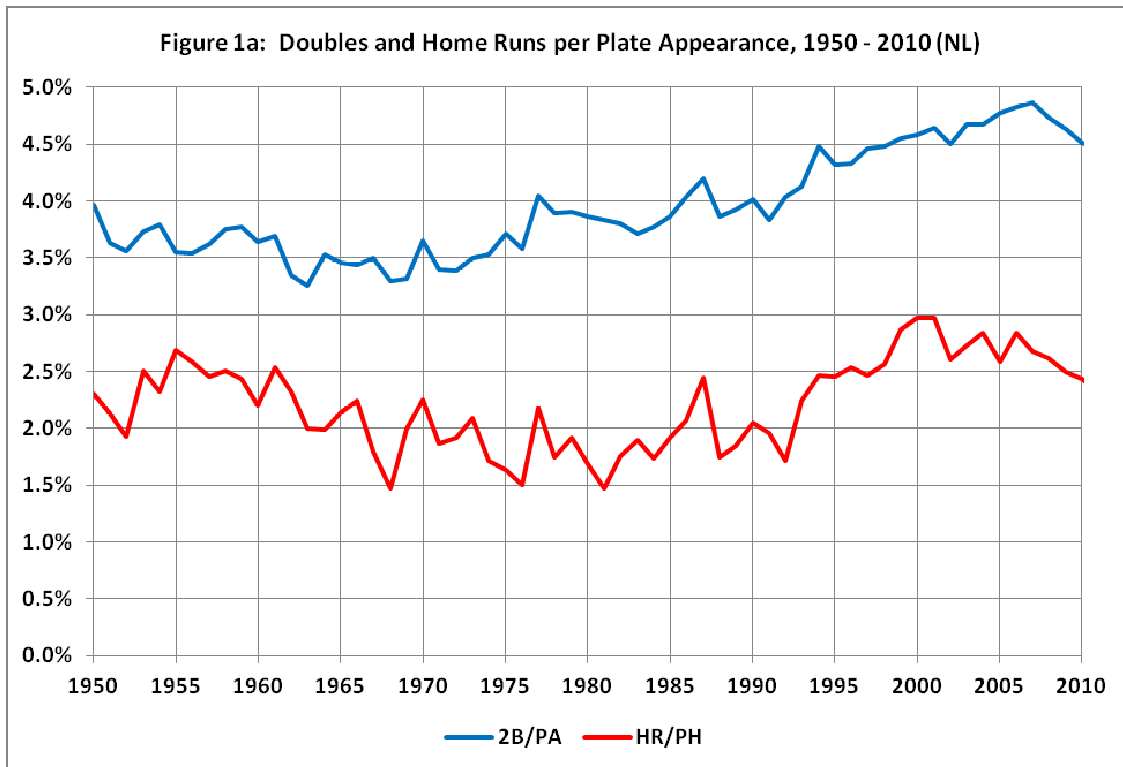
Table 2 shows that in the NL, doubles show the somewhat the same pattern as HRs, with a clear drop-off in the 1980s, and the highest league-leading totals in the 1990s and 2000s. However, the 1970s were not a period of decline in the NL; league-leading doubles totals rose from the 1950s to the 1960s, and again in the 1970s. And the AL doubles pattern is also the same as the AL home-run pattern for the 1980s through the 2000s, with increasing doubles totals in the 1980s compared with the 1970s, and much higher totals in the 1990s and the 2000s. But in the AL, league-leading totals rose, in each decade, from the 1950s to the 1980s.

Table 2 – League-Leading Doubles Totals (Average by Decade)

Decade	NL	AL
1951-1960	40.9	38.6
1961-1970	43.5	39.9
1971-1980	44.1	41.6
1982-1990	41.4	46.7
1991-2000	49.7	49.5
2001-2010	50.0	51.8

Doubles Power: Doubles per Plate Appearance

Both leagues show a fairly consistent, if modest, increase in doubles per plate appearance beginning around 1980 and continuing to the present. (See Figures 1a and 1b.) In both leagues, this upward trend is interrupted from 1988 to 1991, with the decline in doubles per plate appearance being more pronounced in the AL.



³ League-leading totals: http://www.baseball-reference.com/leaders/2B_leagues.shtml. Doubles per plate appearance and HRs per plate appearance calculated from <http://www.baseball-reference.com/leagues/NL/bat.shtml> and <http://www.baseball-reference.com/leagues/AL/bat.shtml>.

Figure 1b: Doubles and Home Runs per Plate Appearance, 1950 - 2010 (AL)

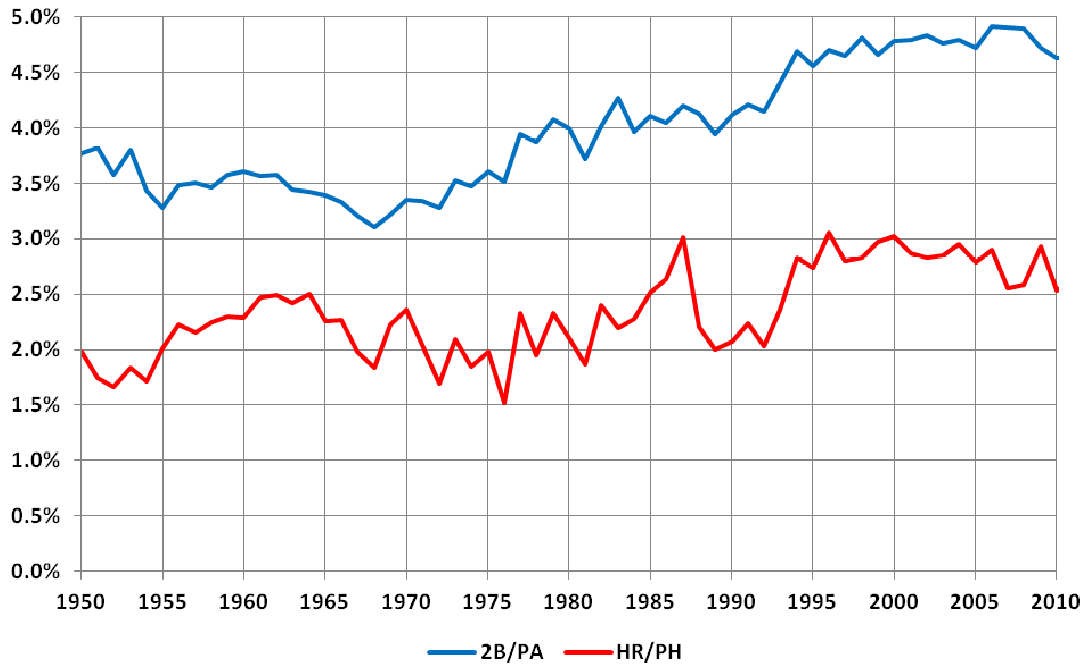
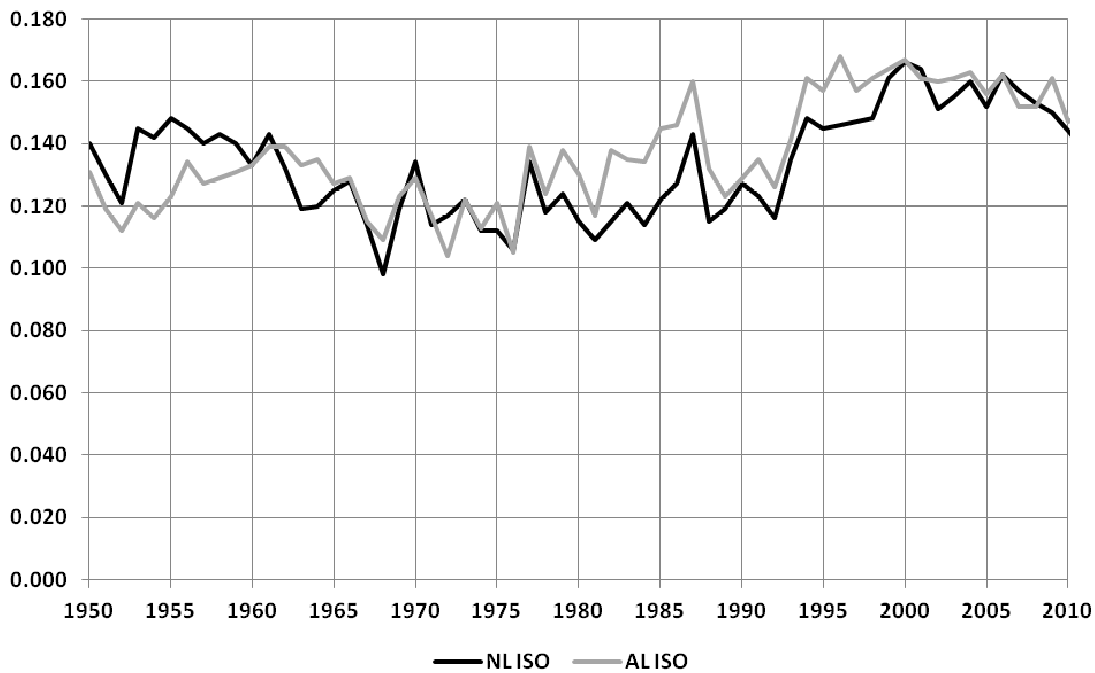


Figure 2: Isolated Power (SA - BA), 1950 - 2010



Home Run Power: HRs per Plate Appearance

The NL saw a consistent increase in HR per plate appearance, from about 1.5% in 1981 to about 2.5% in 1987. (Again, see Figures 1a and 1b.) This was followed by a sharp decline which ended in 1991. In the AL, the pattern was almost identical, albeit starting from a slightly higher level in 1981 (the designated hitter is the most likely explanation for this).

Isolated Power

Figure 2 shows that in both leagues, isolated power began a fairly sharp increase around 1980 (with the AL beginning, as with HRs, at a slightly higher level). There was, again, a sharp decline in the 1988-1991 period, followed by an even sharper rise.

Conclusions

The overall evidence does not, to me, seem to support D'Addona's suggestion that the 1980s were a period of power outage in major league baseball. While league-leading totals in the NL were down from earlier decades (and rose sharply in succeeding decades), league-leading totals in the AL were actually a bit higher than in the 1950s or 1970s (although considerably below the 1960s). In the AL as well, totals shot upward in the 1990s and 2000s.

The evidence from league-leading doubles totals and from doubles and home runs per plate appearance appear to support this conclusion. Once we take account of the dispersion of power among all players, not just at the top, the 1980s do not appear to stand out significantly from the preceding decades.

What we do see, and what we see quite clearly, is that power fell in 1988 and did not rise again until the 1992 season. The "power surge" of 1987, which is clearly evident in Figures 1a, 1b, and 2, apparently led to actions designed to reduce power beginning in 1988. This, not a generalized "power outage," is, it seems to me, the real story of the 1980s.

Donald A. Coffin, dcoffin@iun.edu ♦