By the Numbers

Volume 19, Number 4 The Newsletter of the SABR Statistical Analysis Committee November, 2009

Comment

Book Review: "Beyond Batting Average"

Charlie Pavitt

Lee Panas's new book, "Beyond Batting Average," is a valuable summary of methods for evaluating performance. It's heavily weighted towards material that is available online, at the expense of certain previous research that was published traditionally. However, it remains a good summary of recent sabermetric research in the area of performance evalution statistics.

Lee Panas, Beyond Batting Average: Baseball Statistics for the 21st Century. Available on-line at lulu.com.

Perhaps like many of you, I became aware of this book when I received an announcement courtesy of the author about its publication. I decided to give it a shot if for no other reason than that I am partial to alternative forms of publication, particularly when their upshot is a book that can be downloaded for just seven dollars. I personally found the book to be useful, but that does not necessarily mean that you will also, as its value depends on your current state of knowledge about sabermetrics and the source of that knowledge.

In this issue

I will use Panas's own words to describe the book's intended audience: "My goal is to explain the new world of baseball statistics in a way that any knowledgeable and curious baseball fan will comprehend…it is about how fans can use statistics to enhance their understanding of the game…I have assumed that the reader will be very knowledgeable about baseball and has an interest in looking at new ways to answer questions about the game." Given that goal, and given the fact that it would be unreasonable to expect any one book to do everything, I will attempt to evaluate it within this context. So, for example, we should not expect it to provide anything of note concerning the history of baseball statistics, and it indeed does not, as Panas's use of that history seems to be limited to his reading of John Thorn and Pete Palmer's "*The Hidden Game of Baseball*" and Alan Schwarz's "*The Numbers Game*". Although he does present a brief history, Panas wisely refers interested readers to those sources.

In short, *Beyond Batting Average* is a compendium of methods for evaluating player and team performance from sources almost exclusively available at sabermetric websites. As such, along with fundamental measures such as Runs Created and Linear Weights it includes a number of evaluation indices that have not appeared much if at all in print. This is why I personally found the book valuable, as in my quest to keep up with academic contributions to sabermetrics I don't have the time to do justice to what is online, and I became aware of a number of interesting methods for the first time. Readers who do keep up with online material will not find much of value here unless they just want a summary of such indices in one place.

As such, *Beyond Batting Average* will, indeed, help sabermetric newcomers learn the basics about current evaluation methods. Questions about evaluation are not, however, inclusive of anywhere near the entire set relevant to the game, and Panas's work won't help anyone concerned with other types of issues. Again using Panas's words, after a comment about the impact of the book *Moneyball*, "This is not a book that describes how teams use statistics to manage their organizations." Nor it is a book about in-game strategy, which I personally consider an issue as important as player/team evaluation for the sabermetric neophyte to learn. Again, Panas is savvy enough to send readers to the best sources, *The Hidden Game* and Tango-Litchman-Dolphin's "*The Book"*. And again, I am fine with this, as one book cannot do everything.

My one problem is that limiting one to what is easy to find on-line means one misses important material that is off-line, and leads Panas to make occasional errors. Some are harmless unless, as with academics such as myself, the reader is fanatical about correct attribution. To state two examples: Panas leaves the impression that zone ratings are John Dewan's invention in the late 1980s when, in fact, the idea became obvious the instant that Project Scoresheet data was available and was developed pre-Dewan independently and at about the same time by at least two analysts of which I am aware; and I found it interesting that Lichtman calls his outfielder arm statistic ARM, which is the same acronym that Clem Comly used for his measure of the same skill in *By the Numbers* ten years ago in Volume 10 Number 3.

I understand that none of this is a big deal for the beginning sabermetric student. But ignorance is not bliss when Panas's limitation leads to misleading assertions. It is true, as he noted, that several analysts have concluded that catcher ERA provides no evidence of catcher influence on pitcher performance, but Tom Hanrahan has noted differences based on catcher experience in two studies (*Baseball Analyst* 39 from 1988 and *BTN* Vol. 9 No. 3 from 1999), so there is less unanimity here than the reader might imagine. This, in the end, is my problem with Beyond Batting Average; there is material beyond the Internet world that is, in my view, just as basic as that found on it.

To conclude, if you are not up to snuff on evaluation methods found on-line, you will find this book to be valuable. If you are familiar with them but still want a quick summary, it is probably worth the seven dollars. Just be careful with some of Panas's conclusions.[1]

Charlie Pavitt, chazzq@UDel.Edu ♦

[1] The book is available at http://stores.lulu.com/store.php?fStoreID=873874

"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 .

Right-Wing Insurgency: The Shrinking Advantage of Left-Handed Hitters

Abbott Katz

Over the history of Major League Baseball, left-handed batters, as a group, have always hit for a higher batting average than right-handed batters. However, the gap has been shrinking in recent decades. Here, the author shows us the trend.

Every baseball fan will declare, with almost canonical definitiveness, that theirs is a left-hander's game, its peculiar layout and demographics aiding and abetting the batting averages of swingers from the port side.

And the canon summons two venerable proofs for the left-hitter's edge: 1) the relative closeness of lefties to first base (we'll call it the proximity thesis), and 2) the preponderance of right-handed pitchers in the major leagues – because everybody knows it's easier to hit pitchers dealing with the opposite hand.

And the data in fact are nothing if not supportive of the dictum; considered in the aggregate, left-handed hitters *always* outdo righties, year after year – indeed, *every* year over the past 100. But what's far less evident is the margin's breadth – because the data demonstrate that over that century, the left/right disparity has contracted, and continues to contract.

Organize left-right batting averages into five-year tranches you get something like this[1]:

Years	Average of L	Average of R	Average margin in pct.
1910-1914	.277	.252	10.09%
1915-1919	.270	.245	10.28%
1920-1924	.302	.276	9.27%
1925-1929	.300	.277	8.17%
1930-1934	.296	.271	9.15%
1935-1939	.290	.270	7.21%
1940-1944	.274	.251	9.42%
1945-1949	.273	.254	7.59%
1950-1954	.270	.256	5.60%
1955-1959	.267	.253	5.29%
1960-1964	.260	.250	3.84%
1965-1969	.250	.241	3.84%
1970-1974	.260	.248	4.62%
1975-1979	.269	.256	5.19%
1980-1984	.267	.257	4.01%
1985-1989	.263	.253	3.86%
1990-1994	.268	.257	4.36%
1995-1999	.276	.264	4.53%
2000-2004	.270	.262	2.96%
2005-2009	.267	.263	1.53%

(We might be able to attribute the percentage uptick in 1970-1979 to the rollout of the designated hitter in 1973, with left-handed DHs subbing for feckless right-batting pitchers.)

[1] (Source: Sean Lahman's baseball database, at www.baseball1.com)

How then to reconcile the press towards convergence? Clearly, the proximity thesis serves us not at all here, hinged as it is on a constant – the respective left/right distances to first base, which haven't changed, of course. And you know what they say – you can't explain change with a constant.

On the other hand, one could float the surmise that a boost in the ranks of left-handed pitchers s*hould* depress the left-right margin – i.e., more lefthanded pitchers, lower lefthander batting averages; but in fact the data drift in the opposite direction. Defined by the proportions of innings pitched, it's *right*-handed pitchers who've ratcheted their presence slightly in the last 20 years. So what now?

I would submit that a variation on the Stephen Gould .400-hitter thesis could be impressed into explanatory service here. Gould sought to interpret the present-day dearth in such high-end hitters by pointing to a *general* honing of player competence, having the effect of trimming performance extremes at both ends. Put otherwise, the average hitter has gotten better; and in our case, it would appear right-handed hitters have in large measure managed – eventually – to clear away the impediments posed by the typical, right-hand-pitcher/right-hand-batter matchup. After all, by the time a right-hand hitter cashes his first major league meal money check he's likely spent 15 or so years in the woodshed – in Little League, high school, college, and minor league ball, habituating himself to the right-at-you torque of right-handed pitches. And that apprenticeship is probably far more thoroughgoing and instructive than any tutelage ancient right-hand hitters received. (Of course pitchers have also surely gotten better; batting averages have declined since the 40s, though they've oscillated since. But we're comparing hitters to each *other*, not to pitchers.)

The Gould take is corroborated by the history of switch-hitters, who always position themselves favorably along the left-right pitcher-hitter axis and thus could be expected to bat out higher averages – always. But look at our table, amended:

Years	Average of L	Average of R	Average of Switch
1910-1914	.277	.252	.244
1915-1919	.270	.245	.242
1920-1924	.302	.276	.281
1925-1929	.300	.277	.274
1930-1934	.296	.271	.265
1935-1939	.290	.270	.265
1940-1944	.274	.251	.253
1945-1949	.273	.254	.255
1950-1954	.270	.256	.255
1955-1959	.267	.253	.257
1960-1964	.260	.250	.247
1965-1969	.250	.241	.245
1970-1974	.260	.248	.253
1975-1979	.269	.256	.259
1980-1984	.267	.257	.260
1985-1989	.263	.253	.259
1990-1994	.268	.257	.262
1995-1999	.276	.264	.267
2000-2004	.270	.262	.267
2005-2009	.267	.263	.270
2010	.258	.257	.256

We see switch-hitters at last slinking toward superiority – but only gradually, and only recently. The challenge of fusing major league level left/right batting competence seems to have daunted generations of trial-and-erring forebears, until switch-hitters finally began to get it right.

Of course, we're still left with that pertinacious proximity thesis, which should, after all, continue to enhance left-batters' averages. When the smoke clears, they're *still* closer to first base. But other terms in the batting equation appear to preempt the proximity edge. In a 2007 piece on the web site *The Hardball Times*, John Walsh found[2] that the 2000-2007 base-hit average on ground balls confined to the infield – the kind of rapid-fire event which should surely favor the hitter who has less ground to travel to first – was nevertheless almost identical for left and right-swinging hitters. Walsh reasons that the longer throws of shortstops and third basemen – those fielders more likely to contend with the grounders struck by righthanders – offset the shorter commute lefties need to take to their destination.

[2] http://www.hardballtimes.com/main/article/the-advantage-of-batting-left-handed/

But conjectures notwithstanding, the phenomenon is real and needs to be explained, not explained away. To repeat: the storied advantage of left-handed hitters has given way to a new parity, and so the time may have come to commission a rewrite of the canon. Because today's article of faith is turning into tomorrow's heresy. [3]

Abbott Katz, akatz@hotmail.com ♦

[3] A report on this research appeared in the Wall Street Journal on September 21, 2010, at: http://online.wsj.com/article/SB10001424052748703989304575504023826343654.html