# By the Numbers

Volume 15, Number 4 | The Newsletter of the SABR Statistical Analysis Committee | November, 2005

*Review*

## Freakonomics – The Sabermetrics of Life

Phil Birnbaum

*A review of the recent bestselling book "Freakonomics," which includes studies and methods of interest to sabermetricians.*

*Freakonomics*, one of the best-selling books of 2005, is ostensibly about economics, but its methods and aims are similar to sabermetrics. Further, its authors have blogged about baseball and sabermetrics, which means that *Freakonomics* should be of interest to baseball analysts.

Authors Steven Levitt (a renowned academic economist) and Stephen Dubner (a reporter for the New York Times) begin by pointing out that economics is not strictly about money and GDP, but, rather, about how people respond to incentives. In that vein, the book discusses a number of Levitt's intriguing academic studies, ranging widely in subject area. For instance, Levitt and Dubner tell us that:

- The drop in crime in the early 1990s was caused by the legalization of abortion;

- Drug dealers tend to be "living with their moms" because most of them are at the bottom of a criminal hierarchy and earn much less than minimum wage for their work;

- Children who are read to or taken to museums do not outachieve children who are not; and, in general, a child's achievement is barely influenced by parental actions;

- An analysis of a database of workplaces that sold bagels on the honor system showed that theft by employees was more prevalent during bad weather, at stressful times of year, in larger offices, and among higher-ranked executives;

- Real estate agents manage to get higher prices when selling their own homes than when selling their clients' homes. That's because the agents keep only a couple of percent of any premium price they gain by working hard for their clients' benefit, but they keep all of it when working hard to sell their own house.

Of the studies, the one that's the most obviously sabermetric is the one that shows how Sumo wrestlers cheat.

Sumo wrestlers fight 15 matches. If they finish 8-7 or better, they move up in rank; otherwise, they drop. Therefore, when a wrestler is 7-7, his last match is crucial. But if he's 8-6, the last match is meaningless.

Analyzing data from Japanese sumo matches (obtained from Sumosheet?), Levitt found an interesting anomaly. When a 7-7 wrestler faces an 8-6 opponent, he wins at a .796 clip, even though he has a worse record than his opponent. Indeed, in all other matches where the same two combatants faced each other, the weaker opponent had only a .487 winning percentage.

Is the stronger opponent throwing the match? This evidence alone isn't enough to conclude that. It could simply be that with his ranking on the bubble, the weaker opponent bears down, fights harder, and becomes "clutch," while the opponent, with nothing to lose, avoids overexerting himself.

So Levitt went one step further, and checked what happened

### In this issue

the next time those two wrestlers met. What he found was that in that match, the weaker opponent won only 40% of matches. But one match after that, the percentage went back to the historical average of about 50%.

The conclusion – there seems to be a quid pro quo agreement between the wrestlers, where they agree to exchange one victory for another. Note that each wrestler benefits, winning a more-important match in exchange for throwing a less-important one. That is, when there is an opportunity to profit from a change in behavior, the wrestlers respond to incentives.

If you're still not convinced, the authors have additional evidence. Levitt found patterns of reciprocation not just between individual wrestlers, but between "stables" (groups of wrestlers with a common manager who do not compete against each other). More devastating, the patterns disappear completely during times when there are strong allegations of match-rigging. When the heat is on, the tendency of bubble wrestlers to outperform vanishes completely, with the winning percentages dropping back to the expected .500.

And, if that's not enough, there's a topper: a few years ago, two sumo wrestlers came forward, Canseco-style, with specific allegations of wrongdoing. They named names, both corrupt wrestlers and honest ones. Those cited as honest showed no pattern of clutchness, either for or against. The allegedly dishonest ones did.

This study is almost pure sabermetrics. And, indeed, a fair bit of sabermetrics would qualify as economics. The famous question of whether players outperform in the last year of their contract is almost pure Freakonomics. There's the issue of clutch hitting – do batters "turn it on" when they have stronger incentives to succeed? And there have been studies on whether National League pitchers are less likely to hit opposing batters than their American League counterparts, since they have to bat and so might face a personal plunking in return. All these questions would be well at home in Levitt and Dubner's book.[1]

**Freakonomics**

**By Steven D. Levitt and Stephen J. Dubner**

**Morrow, 242 pages, $25.95 (US)**
**ISBN 006073132X**

And while the other studies in the book aren't sports related, they do have the flavor of sabermetric studies – which, I guess, is inevitable in any field that mines databases for evidence of relationships in human endeavors.

Do children given "black" names like DeShawn and Roshanda do worse in life than black children with "whiter" names? After adjusting for league and park – er, I mean, factors such as education and other social and economic influences – Levitt says the answer is no. The question reminded me of Bill James' study of draft choices, and whether college picks turned out better than than players drafted out of high school; that study would fit right into *Freakonomics*.

Or, perhaps not. One thing about the book that frustrated me a bit was that we're not given much information about the study or the evidence itself – just the conclusion. In a mass-market book, that isn't necessarily a bad thing. But readers who have been spoiled by the Bill James approach – "here's the study, here's why I did it this way, here's what I found, here's my conclusion, here are reasons my conclusions might be wrong, and here's the data if you want to check or extend the study" – might feel a little shortchanged.

The chapter on baby names, for instance, doesn't give any of the details of just *how* Levitt found out that black names don't matter much. Instead, the authors go on to show the "whitest" and "blackest" baby names[2], and to explain how names spread from the upper classes to the lower classes.

In that chapter, Levitt and Dubner present, as fact, a boy named "LemonJello" (but pronounced "le-MONZH-ello"). Baby name guru Laura Wattenberg, in her blog[3], argues that the story is an urban legend. (For the record, neither party mentions former pitcher Mark Lemongello.) Wattenberg also points out other shortcomings in the chapter's analysis, and accuses the authors of walking into an unfamiliar area of knowledge without realizing that their "revelations" have already been improved upon by others in the field. Levitt and Dubner defend themselves on their own blog[4], but I didn't find their defense totally convincing. In any case, I suppose the point is that even if you're able to make insightful discoveries from a set of data outside your own field, that doesn't turn you into an expert in that field.

---

[1] Just as this issue was going to press, I found that Levitt actually published an academic article on the plunking issue – the study is almost pure sabermetrics, and would fit right into BTN. It's at http://pricetheory.uchicago.edu/levitt/Papers/LevittHazardsOfMoralHazard1998.pdf .

[2] The top three in each race/sex group: Jake, Connor and Tanner; Molly, Amy, and Claire; DeShawn, DeAndre, and Marquis; and Imani, Ebony, and Shanice.

[3] Wattenberg's comments are at http://www.babynamewizard.com/blog/2005/04/name-onomics.html and http://www.babynamewizard.com/blog/2005/05/names-race-and-economists.html .

[4] The authors' blog is at www.freakonomics.com/blog . Their response to Wattenberg can be found by searching for her name.

Which brings us back to sabermetrics.

There's nothing about baseball in *Freakonomics*. But in the authors' blog, Levitt talks a bit about baseball. One of his themes is that he is skeptical of the *Moneyball* story of the A's strategy. Specifically, he doesn't believe the A's success was indeed due to exploiting the market for low-priced OBP. For five teams, he lists their average offensive line from 2000 to 2004:

| Team | R | HR | BB | K | BA | OBP | SLG | OPS |
|---|---|---|---|---|---|---|---|---|
| Team A | 867 | 200 | 591 | 1045 | .276 | .348 | .454 | .802 |
| Team B | 865 | 222 | 638 | 1022 | .271 | .351 | .450 | .801 |
| Team C | 838 | 202 | 633 | 1029 | .264 | .343 | .436 | .778 |
| Team D | 829 | 193 | 575 | 1041 | .269 | .341 | .437 | .778 |
| Team E | 828 | 159 | 619 | 1022 | .275 | .349 | .422 | .771 |

Levitt asks, which is the A's? And doesn't it look like all five teams are generating runs in exactly the same way?

It's a reasonable argument, but I'm not sure I agree with it. The teams scored about the same number of runs, but, in baseball terms, are the lines really similar? The top team in batting average hit .276 – that would have been second in the AL in 2005. The bottom team hit .264, which would have been tenth. That's quite a difference, second to tenth. And I haven't checked, but I'd bet that if you averaged five years of the American League, you'd reduce the variance so much that the difference between .276 and .264 would be a lot more than eight places in the standings.

The A's are Team C. They had the lowest batting average and the second-highest walks. They finished easily in the top half of the league in runs. (And all this on a mid-market team's budget.) I think that's reasonably consistent with the idea that the A's are doing things the way *Moneyball* claimed they are.

Levitt has other blog posts on the A's and sabermetrics – one involves the probability of the 2005 Royals' 19-game losing streak. They are all worth reading, as are many of the responses from sabermetrically-literate readers.[5]

And there may be more to come. Significantly, Levitt writes in his blog that he will "debunk some sabermetric ideas" in the future. There's been nothing yet, but I'm looking forward to it.

*Phil Birnbaum, birnbaum@sympatico.ca* ♦

[5] To find these postings, head to www.freakonomics.com/blog, and search for "Beane" or "Royals" or "A's."

## Informal Peer Review

The following committee members have volunteered to be contacted by other members for informal peer review of articles.

Please contact any of our volunteers on an as-needed basis - that is, if you want someone to look over your manuscript in advance, these people are willing. Of course, I'll be doing a bit of that too, but, as much as I'd like to, I don't have time to contact every contributor with detailed comments on their work. (I will get back to you on more serious issues, like if I don't understand part of your method or results.)

If you'd like to be added to the list, send your name, e-mail address, and areas of expertise (don't worry if you don't have any - I certainly don't), and you'll see your name in print next issue.

Expertise in "Statistics" below means "real" statistics, as opposed to baseball statistics - confidence intervals, testing, sampling, and so on.

| Member | E-mail | Expertise |
|---|---|---|
| Ben Baumer | bbaumer@nymets.com | Statistics |
| Jim Box | jim.box@duke.edu | Statistics |
| Keith Carlson | kcsqrd@charter.net | General |
| Dan Evans | devans@seattlemariners.com | General |
| Rob Fabrizzio | rfabrizzio@bigfoot.com | Statistics |
| Larry Grasso | l.grasso@juno.com | Statistics |
| Tom Hanrahan | Han60Man@aol.com | Statistics |
| John Heer | jheer@walterhav.com | Proofreading |
| Dan Heisman | danheisman@comcast.net | General |
| Bill Johnson | firebee02@hotmail.com | Statistics |
| Mark E. Johnson | maejohns@yahoo.com | General |
| David Kaplan | dkaplan@UDel.Edu | Statistics (regression) |
| Keith Karcher | karcherk@earthlink.net | Statistics |
| Chris Leach | chrisleach@yahoo.com | General |
| Chris Long | clong@padres.com | Statistics |
| John Matthew IV | john.matthew@rogers.com | Apostrophes |
| Nicholas Miceli | nsmiceli@yahoo.com | Statistics |
| Duke Rankin | RankinD@montevallo.edu | Statistics |
| John Stryker | johns@mcfeely.interaccess.com | General |
| Tom Thress | TomThress@aol.com | Statistics (regression) |
| Joel Tscherne | Joel@tscherne.org | General |
| Dick Unruh | runruhjr@dtgnet.com | Proofreading |
| Steve Wang | scwang@fas.harvard.edu | Statistics |

# World Series Hitting by Pitchers and Designated Hitters

Ray Ciccolella

*Conventional wisdom has it that in the World Series, the National League pitchers should outhit those of the AL, because they have more experience. And, the reverse should be true for DHs. Here, the author examines some 30 years worth of World Series games to see if that is indeed the case.*

Each year as the Fall Classic is analyzed there is discussion about how the American League team is at a disadvantage because of the designated hitter (DH). The standard thinking is that in those games where the pitchers bat, the National League benefits, since their pitchers bat throughout the season and therefore can be expected to outperform their American League counterparts.

Does such an advantage actually exist? If so, how big is this advantage? Also, does the American League have a potential advantage in their home games since they have players who are more accustomed to being the DH?

To address these questions I examined the World Series batting performance of the designated hitters and pitchers for both leagues from 1976 (the first year a DH was used in the Series) through 2005. I only had access to summary box scores, so I focused on the statistics available to me: At Bats, Hits, Runs, Runs Batted In, and Batting Average. I also created an index I called Runs Produced per Game which I defined as:

$$RP/G = \left(\frac{R + RBI}{2}\right) \text{ per 27 hitless at-bats}$$

Tables 1 and 2 show the batting performance of pitchers and designated hitters in the World Series since 1976 for each league. I also included overall league performance in the Series in the bottom row of each table.

**Table 1 – American League Performance in World Series, 1976-2005**

| | At Bats | Hits | Runs | RBI | Batting Avg | Outs | RP/G |
|---|---|---|---|---|---|---|---|
| DH | 297 | 68 | 42 | 38 | .229 | 229 | 4.72 |
| Pitcher | 193 | 12 | 6 | 7 | .062 | 181 | 0.96 |
| DH + P | 490 | 80 | 48 | 45 | .163 | 410 | 3.06 |
| AL Total | 5617 | 1445 | 712 | 681 | .257 | 4172 | 4.51 |

**Table 2 – National League Performance in World Series,1976-2005**

| | At Bats | Hits | Runs | RBI | Batting Avg | Outs | RP/G |
|---|---|---|---|---|---|---|---|
| DH | 298 | 74 | 37 | 34 | .248 | 224 | 4.28 |
| Pitcher | 166 | 20 | 11 | 6 | .120 | 146 | 1.57 |
| DH + P | 464 | 94 | 48 | 40 | .203 | 370 | 3.21 |
| NL Total | 5588 | 1414 | 683 | 652 | .253 | 4174 | 4.32 |

It is clear that the NL pitchers, as expected, have outperformed the AL pitchers in the Series. Their batting average is almost double and their run production per game is more than 50% higher. The American League DHs did outperform the NL DHs from a runs produced perspective, both on a total and per game basis, despite a lower batting average.

Combining the DHs batting performance with those of the pitchers, the NL comes out ahead. The American League does appear to have an advantage of 5 extra RBI, but those were obtained with additional outs.

The higher number of outs is partially attributed to AL pitchers getting 27 more at bats than NL pitchers. One reason AL pitchers have more at bats is because the AL has a 92-73 record in the 165 World Series games since 1976. Pitchers get more at bats on winning teams because those teams pitch hit for their pitcher less frequently. The lower batting averages for the American League account for the rest of the extra outs.

The AL pitchers accounted for 6.7 games at 27 outs per game and produced 0.63 (1.57 less 0.96) fewer runs per game than NL pitchers. These figures translate to a difference of 4.2 runs (6.7 multiplied by 0.63).

When I include the performance of the designated hitters, the difference drops. The AL pitchers-plus-DHs were responsible for 410 outs, or 15.2 games. At a difference of 0.15 runs per game (3.21 less 3.06), that translates to only 2.3 extra runs.

## Regular Season Performance

While for the World Series games I did not have access to more detailed batting statistics I was able to find batting average and on-base plus slugging (OPS) data by league and position on the Baseball Prospectus website. Table 3 covers 1997-2005, the period of regular season interleague games.

The pattern from the World Series repeats in the regular season. National League pitchers outhit the AL pitchers in the regular season every year since 1997 in terms of batting average and OPS. Similarly, we can also see that AL DHs have out-performed their NL counterparts overall on both measures, though not in every season.

As we would expect because of the reduced number of at-bats, the performance of the AL pitchers and NL DHs vary more from year to year than the performance of the AL DHs and NL pitchers. The performance of the latter two is very consistent from season to season.

**Table 3 – Regular Season Batting, 1997-2005**

| | AL Pitchers | | NL Pitchers | | AL DHs | | NL DHs | |
|---|---|---|---|---|---|---|---|---|
| Season | BA | OPS | BA | OPS | BA | OPS | BA | OPS |
| 1997 | .114 | .293 | .139 | .336 | .274 | .795 | .261 | .741 |
| 1998 | .126 | .306 | .147 | .354 | .277 | .825 | .221 | .624 |
| 1999 | .132 | .349 | .148 | .358 | .278 | .840 | .255 | .738 |
| 2000 | .117 | .290 | .150 | .363 | .276 | .823 | .288 | .794 |
| 2001 | .136 | .302 | .144 | .343 | .260 | .786 | .260 | .751 |
| 2002 | .136 | .311 | .148 | .356 | .264 | .788 | .228 | .733 |
| 2003 | .128 | .296 | .144 | .347 | .260 | .787 | .293 | .879 |
| 2004 | .092 | .237 | .146 | .348 | .263 | .789 | .245 | .761 |
| 2005 | .118 | .304 | .150 | .354 | .259 | .777 | .249 | .714 |
| Avg | .122 | .299 | .146 | .351 | .268 | .801 | .256 | .748 |

National League pitchers have an OPS .052 higher than AL pitchers, almost exactly the .053 difference for AL over NL DH's. Using a quick correlation of OPS to runs scored, a .052 advantage in OPS is worth about 0.60 runs per game. This estimated difference in runs per game is similar to, but higher than, the difference in World Series runs produced per game between NL and AL pitchers, 1.57 versus 0.96, and AL and NL designated hitters, 4.72 versus 4.28 (see Tables 1 and 2).

While this data still represents a limited number of games for NL DH's and AL pitchers, it nonetheless represents about 12 times as many games as the World Series data. Also this data is not completely clean, as the at bats for the AL pitchers and NL DH's all took place in inter-league play while the at bats for NL pitchers and AL DH's took place across all games. The number of years involved and the consistency of the results, however, gives me reasonable confidence that the differences in performance are real despite these two potential data issues.

## Conclusion

The conventional wisdom is at least partially correct. In the World Series the NL pitchers out hit the AL pitchers by a reasonable margin. The AL has had an advantage with their designated hitters, which closes the gap somewhat.

While the results might be statistically significant, however, I do not believe they are significant in the practical sense. From 1976 through 2005 there have been 165 World Series games played, about one season's worth. Considering just the pitchers my estimated difference is about 4 runs over those games. The difference drops to about 2 runs once I include the DH's. Those 2 runs are worth about ¼ an extra win over 29 series – hardly noticeable.

My conclusions are based on less than 500 at bats per league in the World Series and somewhat limited measures of performance. I have greater confidence in my findings and conclusions about World Series batting performance, however, since the patterns displayed in Series games are very similar to those seen in regular season play.

*Ray Ciccolella, rciccolella@austin.rr.com* ♦

## 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, either by e-mail or on CD. 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).

If your submission discusses a previous BTN article, the author of that article may be asked to reply briefly in the same issue in which your letter or article appears.

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
88 Westpointe Cres., Nepean, ON, Canada, K2G 5Y8
birnbaum@sympatico.ca

## Announcement: Baseball, Statistics, and the Role of Chance in the Game

The Ohio Section of the Mathematical Association Association is sponsoring a short course on "Baseball, Statistics, and the Role of Chance in the Game" at Mt. Union College, Alliance, Ohio on July 7-9, 2006.

Jim Albert, Professor of Statistics at Bowling Green State University, will introduce several explorations of baseball data including the search for the ultimate batting statistic, looking for true streakiness and clutch ability, modeling run production by a Markov Chain model, and comparing great players such as Babe Ruth and Barry Bonds who played in different eras. This workshop will be directed both to instructors who wish to infuse their teaching of probability and statistics with applications from baseball and to baseball fans interested in learning about sabermetrics.

For more information, visit the meeting website at www.muc.edu/~zwilliml/shortcourse .

# Finding Implicit Linear Weights in Run Estimators

Brandon Heipp

*The structure of the Linear Weights formula is simple enough that we can immediately observe the marginal values of events – that, for instance, a single is worth 0.47 runs. But what is a single worth in multiplicative formulas like Runs Created or Base Runs? Here, the author shows us how to figure this more difficult calculation.*

Run estimators used in sabermetrics can generally be divided into two classes—linear methods and what I will call for lack of a better term multiplicative methods. Linear methods, like Pete Palmer's Batting Runs or Paul Johnson's Estimated Runs Produced, assign each offensive event (such as singles, walks, and outs) with a coefficient that reflects the average number of runs that the event produces. These weights are sometimes found through empirical means (play-by-play and base/out table analysis), experimentation, or multiple regression. These methods, when summed for all players on a team, will yield the same result as if the estimator was applied to the team totals.

Multiplicative methods, exemplified by Bill James' Runs Created and David Smyth's Base Runs, attempt to model the team scoring process. James does this by multiplying an "on base" factor by an "advancement" factor and dividing by an "opportunity" factor, while Smyth's construct multiplies baserunners by the proportion of baserunners who score, and then adds home runs. In both cases, multiplication and division are used, and an event does not have a specified run value. If the estimates are summed for players on a team, the total will usually be close but not equal to the team estimate.

Because the two classes of run estimators use different approaches, it is difficult to compare the emphasis that each puts on various events. We know that Batting Runs pegs a single at .47 runs, but what weight does RC put on a single? While the formula does not explicitly tell us, we can use mathematical approaches to get an answer.

Suppose we use Bill James' most elementary RC formula, RC = (H+W)*TB/(AB+W), and apply it to the 1979 Pirates. The Pirates offensive statistics were:

| AB | H | 2B | 3B | HR | BB |
|---|---|---|---|---|---|
| 5661 | 1541 | 264 | 52 | 148 | 483 |

Therefore their RC estimate was 775.142. To approximate the value of a single to them, we can add one single to the totals and recalculate RC. One more single means one additional AB, one additional hit, and one additional total base. These new totals give an RC estimate of 775.728, an increase of .586. Therefore, the approximate value of a single to the team, according to RC, was .586 runs.

However, when we add a single, we change ever-so-slightly the context. Suppose we added 100 singles. The Pirates' team BA would jump from .272 to .285, a substantial change. The change in context would affect the events of all values in a multiplicative methods, and the change in RC from 100 singles divided by 100 would be .593, compared to .586 from adding one. While adding one single only slightly changes the context, it does to some extent.

If we want to isolate the value of a single in the exact context of the 1979 Pirates, we need to add as small of an amount as possible of each event, then recalculate RC. We should eventually be able to add an infinitesimal amount that will allow us to find the precise value of a single according to RC without changing the context at all. In calculus terms, we want to find the limit as x approaches zero of the RC with x additional singles minus the original RC, divided by x. It turns out that we can find this by taking the partial derivative of the RC equation with respect to singles.

I will skip the details of the calculus (partly because am not qualified to explain it), and just show the final result. Let A, B, and C be the actual team totals in the A, B, and C factors of Runs Created and let a, b, and c be the coefficient of a given event in the A, B, or C factor. Then the linear weight of that event according to RC is:

$$LW = \frac{ACb + BCa - ABc}{C^2} = b\frac{A}{C} + a\frac{B}{C} - c\frac{A}{C}\frac{B}{C} \qquad (1)$$

In the case of a single, a, b, and c are all equal to 1 (since each single counts as 1 hit, 1 total base, and 1 plate appearance – that is, adding a single increases A, B, and C each by 1. So, for a single, the above formula reduces to:

$$LW_{1b} = \frac{AC + BC - AB}{C^2} = \frac{A}{C} + \frac{B}{C} - \frac{A}{C}\frac{B}{C} \qquad \text{(1a)}$$

For a home run, the A and C factor increase by 1, but the B factor (total bases) increases by 4. Therefore a=1, b=1, and c=4, so:

$$LW_{HR} = \frac{4AC + BC - AB}{C^2} = 4\frac{A}{C} + \frac{B}{C} - \frac{A}{C}\frac{B}{C} \qquad \text{(1b)}$$

In the case of the Pirates, their totals were:

```
A   = H + W = 1541 + 483 = 2024
B   = TB = 1541 + 263 + 2*52 + 3*148 = 2352
C   = AB + W = 5661 + 483 = 6144
A/C = 2024/6144 = .329427
B/C = 2352/6144 = .382813
```

And so using formulas (1a) and (1b), respectively,

```
LW(single)   = .329427     + .382813 - .329427*.382813 =  .58613
LW(home run) = .329427*(4) + .382813 - .329427*.382813 = 1.57441
```

In the same way we can find the LW for any event (the LW of an out by this method will be the absolute value like that presented in ERP, not the average value as seen in Batting Runs). The derivative formula above can be applied to any variant of RC based on the A*B/C construct.

David Smyth's Base Runs is an alternative multiplicative method that I previously profiled in the August, 2001 BTN. It has superior accuracy to RC outside of the normal range of major league teams and is just as accurate of an estimator of team runs scored. Smyth defined BsR as A*B/(B+C) + D, where A is baserunners, B is advancement, C is outs, and D is home runs. B/(B+C) is an empirical estimate of the percentage of baserunners that will score.

Smyth's most basic BsR equation is:

```
A = H + W – HR
B = 1.428*TB - .612*H – 3.06*HR + .102*W
C = AB – H
D = HR
```

The derivative formula is more complicated, but defining terms similarly to above, it is:

$$LW = \frac{(B+C)(Ab+Ba) - AB(b+c)}{(B+C)^2} + d \qquad \text{(2)}$$

The 1979 Pirates have an A factor of 1876, a B factor of 2013.378, and a C factor of 4120. Suppose we want to find the implied LW value of a triple from BsR. A triple has an A coefficient of 1 (a = 1), a B coefficient of 1.428*(3)-.612*(1) = 3.672 (b = 3.672), and C and D coefficients of zero (c = 0, d = 0). Then we plug these values into equation (2):

$$LW_{3b} = \frac{(2013.378 + 4120)(1876 * 3.672 + 3013.378 * 1) - 1876 * 2013.378(3.672 + 0)}{(2013.378 + 4120)^2} + 0 = 1.08272$$

At this point you may be saying, "All that calculus is nice, but what is it useful for?" There are any number of possible uses for the estimated linear weights, and I will discuss a few of them.

One is to find approximate linear weight values for a team or league. Empirical LW, calculated by using play-by-play data, are unavailable for many seasons because the play-by-play data does not exist (although if it doesn't, Retrosheet is assuredly working on it). Using a formula like BsR that we are confident will give reasonable results in a variety of contexts, we can estimate the LW values for, say, the 1949 National League, or the 1961 Yankees.

Another possible use is to calculate player value in different contexts. We can estimate Alex Rodriguez' runs contributed by using the linear weights of his actual team. Or we could estimate how many runs his 2005 statistics would have been worth in the context of a different team (say, the 2005 Red Sox), or another league (assuming that he did not substantially change the context, which is probably a faulty assumption for a team and a fairly safe assumption for a league).

A third possible use is to use the method to hone the coefficients used in the BsR or RC equation. If we accept that the empirical LW are a good estimate for the value of each event, we can rearrange the derivative formula to solve for the B coefficient of each event. (We would focus on the B coefficient because each event has a fairly clear coefficient in the A, C, and D factors. It is the advancement coefficients that require the most estimation to find.) If L is equal to the empirical LW of an event, then this equation returns B' ("B prime"), the B coefficient of the event in that context that is necessary to return an estimated LW equal to L:

$$b = \frac{(B'+C)^2(L-d) - a(B'^2 + B'C) + AB'c}{AC} \qquad (3)$$

Here I have defined B' to be the B value necessary to make the BsR estimate equal to the actual number of runs scored. We cannot use the calculated B value because it will change based on the different B coefficients we find for each event. But if we know the empirical LW values, then applying them should result in the exact total of runs scored. So if the actual number of Runs Scored is R, and A, C, and D are known but the exact B is not, then that Exact B (B') can be found thusly:

$$B' = \frac{C(R-D)}{A-R+D} \qquad (4)$$

For example, suppose we knew that the empirical value of a triple for the 1979 Pirates was 1.02 runs (I pulled this number out of thin air for the purposes of example; it is not the actual value). Since we know the Pirates scored 775 runs, we can plug numbers into equation (4) and obtain that their B' is:

$$B' = \frac{4120(775-148)}{1876-775+148} = 2068.247$$

Then we can find the needed b coefficient for a triple using equation (3):

$$b = \frac{(2068.247+4120)^2(1.02-0) - 1(2068.427^2 + 2068.427*4120) + 1876*2068.247*0}{1876*4120} = 3.398$$

If we had the empirical LW for each offensive event, we could apply this formula to all of them and get B coefficients for Base Runs that resulted in a perfect estimate of runs scored given the actual totals of that team. So if we have confidence that the empirical LW values accurately reflect the value of each event, we can use them to ensure that our multiplicative method will return the same values.

To use this approach with Runs Created, you must calculate B' as follows:

$$B' = \frac{R*C}{A} \qquad (5)$$

Then the b value for the given event in RC is given by:

$$b = \frac{LC^2 - BCa + ABc}{AC} \tag{6}$$

I'm sure that other sabermetricians can come up with other possible uses for this approach. If nothing else, it allows one to see how a multiplicative method has implicitly weighted each offensive event in a particular context, and thereby provide insight into why the runs scored estimates of the linear and multiplicative methods may diverge.

## References

This post on the FanHome "Strategy and Sabermetrics" message board by David Smyth describes the Base Runs method: http://mb7.scout.com/fbaseballfrm8.showMessage?topicID=1045.topic

This series of articles by Tango Tiger discusses run estimation methods in general and includes detailed discussion of empirical Linear Weights and Base Runs: http://tangotiger.net/#Baseruns

This article by Kevin Harlow applies the partial derivative approach to Base Runs as well: http://members.cox.net/~harlowk22/br1.html

*Brandon Heipp, bcheipp@yahoo.com* ♦

# More on OBP vs. SLG

Mark Pankin

*In a previous BTN, an article by Phil Birnbaum presented a method for determining how much a point of OBP is worth (in terms of run scoring) compared to a point of SLG. In this, one of two studies this issue on the topic, the author presents a different method and obtains a significantly different result.*

Most of us were thrilled when we heard about and read *Moneyball* by Michael Lewis. One reason is that it discussed many topics that we could research further. Perhaps the most "famous" was the assertion on page 128 that Paul DePodesta, who was Billy Beane's assistant at the time and is now the former GM of the Dodgers, modified the Runs Created formula to make it more accurate and used his version to find that "an extra percentage point of on-base percentage was worth three times as much as an extra percentage point of slugging percentage." Phil Birnbaum wrote about that in the May 2005 issue of *BTN*, and I and others have also analyzed and reported on this issue. In this article I discuss my research, which I presented at SABR in Cincinnati in 2004.[1]

The popular statistic OPS—on-base percentage (OBP) plus slugging average (SLG)—implies that an extra point of either quantity is equally valuable (with respect to run scoring). However, all analyses conclude that additional OBP is worth more than the same increase in SLG. That should not be a surprise because the reason a team does not score or stops scoring in an inning is not failure to hit for extra bases, but because it has made three outs. Although the importance of OBP skills has gained widespread acceptance only fairly recently, some were aware of it long ago. Earl Weaver in the 1984 book *Weaver on Strategy* (with Terry Pluto) states his fourth "law" on page 39 as "your most precious possessions on offense are your twenty-seven outs." That emphasizes the importance of not making outs and the benefits of a high on-base percentage. Due to its simplicity and now widespread acceptance, I don't think OPS is going away anytime soon, and it is better than batting average or HR and RBI. Later I will present an "improved" version of OPS.

In his article, Phil found that OBP was just a little more valuable than SLG and found in his calculations, based on the 1987 AL, that an extra point of OBP was worth 1.2 extra SLG points. I'll refer to this quantity as the marginal value ratio (MVR) in the rest of this article. Phil and I have exchanged e-mails about the calculation methods used, and I think Phil has increased his estimate of the MVR based on revised calculations. I hope he will either add some comments to this article or write a follow-up article for this issue.

A more typical estimate of MVR is around 2, and in a bit I will discuss how I came to that number. I don't think there is any "true" MVR as the run scoring effects depend on the rest of the team and the calculation method used. As I discuss in a bit, the position in the lineup has a dramatic effect on the MVR of a particular player batting in a particular batting order position in the team's lineup. I found MVRs ranging from roughly from 2 to 3 for the 2001 Oakland team.

Assuming that 2, which is a "nice" whole number, is a reasonably accurate value for general use an player evaluation, we can use the following as an improved OPS:

```
OBP + SLG + (OBP – 0.340)
```

I chose 0.340 because it is a typical league average OBP in recent years. In effect the formula gives a player a "bonus" for each point above the league average OBP and a penalty for each point below. Subtracting 0.340 does not affect player comparisons and is in a sense irrelevant unless we change the value for different eras or perhaps different league-seasons. Since I doubt my improved version is going to get widespread use, I am not going to analyze that aspect of it further.

My Cincinnati presentation slides and a few notes are on my website.[2] There are additional topics in the presentation and more details for the calculations than what follows here.

---

[1] Alan Schwarz, author of the 2004 book *The Numbers Game***,** told me at that SABR meeting that Lewis misinterpreted DePodesta's finding as we have read it. Instead of applying to scoring runs, three is the relative additional runs creation from investing in additional OBP as compared to trying to obtain additional SLG. In other words, it costs three times a much to buy additional scoring by increasing SLG as it does by increasing OBP. I have no idea if this is so, and I have not tried to get in touch with DePodesta.

[2] http://www.pankin.com/sabr34.pdf

My basic approach to determining the MVR is similar to that used by Phil and others. The idea is first to change either OBP or SLG by a certain amount, determine the effect of that change on run scoring. Then one can either figure out

(a) what change in the other is needed to have the same effect on scoring, or
(b) the effect on scoring of the same change in the other quantity.

I used approach (a), but (b) is equally valid and should produce a similar MVR. There are a couple of considerations to doing this. One is what to vary in order to change OBP and SLG. The other is how to estimate the effects on runs scored. Also, one or more baseline cases, such as the 1987 AL that Phil used, need to be chosen for the analysis.

To vary OBP and SLG I took a simple approach that has the advantage of making it possible to change one of OBP or SLG without changing the other. By changing the number of walks and leaving other values alone, OBP can be changed without affecting SLG. By changing the distribution of hits (among singles, doubles, triples, and homers), but not total hits, SLG can be varied without affecting OBP. This approach likely is not "realistic" in that a player or total team is unlikely to have such a pure change. However, determining what is realistic adds complications, and it is not obvious that a more realistic approach will lead to a more accurate estimate of MVR. Phil took a different approach than I did and made changes that affected both quantities at the same time. These types of differences are likely to result in different MVR estimates, but ones that should not be too large if the same baseline is used.

To estimate the effects on scoring of changes in OBP and SLG, I used two techniques. One was a recent version of the Runs Created formula. The other was applying my Markov model, which I have discussed and written about quite frequently. More details, probably more than you want, can be found at my website.[3]

Because *Moneyball* focused on the 2001 A's, I decided to use that team as a baseline. To see if they differed from the leagues, I also did the calculations for the 2001 AL and NL. My approach was to change the number of walks to produce increases and

## Phil Birnbaum Responds

Thanks to Mark for the opportunity to add some comments.

First, Mark's results will differ slightly from mine because Mark used walks (only) to increase OBP, while I used a combination of hits and walks. The OBP/SLG ratio is higher for walks than for hits because a walk adds only to OBP, while a hit adds to both OBP and SLG. One way of thinking about this is that a hit adds to SLG to acknowledge the advancement value of the hit. But the walk, also, has an advancement value (although a smaller one), but makes no contribution to SLG. If a walk counted for the appropriate amount of "bases" in SLG – say, half a base – the ratios would be identical. But since it does not, the walk increases the OBP/SLG ratio more than the hit.

If I had used only walks to increase OBP, as Mark does, I would have come up with a ratio of 1.37 instead of 1.2. This is still significantly different from Mark's finding.

Secondly, and as Mark mentions, I have found that my calculation should be revised. The ratio between OBP and SLG is very sensitive to the coefficients used in the linear weights formula. In my original study, I used Pete Palmer's standard weights. But I came across a previous analysis I did of play-by-play data for the 1988 American League, and found that the empirically-observed weights were significantly different. Specifically, a single was .465 (instead of .46), an extra base was .329 (instead of .35), a walk was .302 (instead of .33), and an out (calculated to zero out the league) was -.28072 (instead of -.25).

Repeating the analysis in my original article, but with the new weights, gives a new ratio of 1.43. Even considering walks alone, the new ratio is 1.5. Clearly, 1.5 is still significantly below Mark's figure of approximately 2.

Finally, the only other reason I can think of for the discrepancy is the seasons used. Mark used the 2001 Oakland A's, which had a signficantly higher offense than the 1987 and 1988 American League on which my study was based. As offense increases, the ratio should also increase. To see why, imagine a team that rarely makes an out. In that case, extra bases are almost worthless, since anyone who reaches base is likely to score regardless; but getting on base is almost always worth a run. The ratio should therefore tend towards infinity as OBP approaches 1.000.

I'm not sure if this effect is sufficient to account for the difference between our results, but it probably accounts for at least a small part of it.

[3] http://www.pankin.com/baseball.htm

decreases of 10 and 20 (0.010, 0.020) OBP points and then calculate the expected change in runs for the season in each case. To vary SLG without changing OBP, I decided to increase or decrease the extra base hits to maintain the same proportions among the doubles, triples, and homers as the actual full season totals. The differences were subtracted from singles when SLG was increased or added to singles when SLG was decreased. In all cases, the numbers of hits, at bats, and walks were kept the same so OBP was not affected. SLG was varied by the amounts needed to produce the same changes in runs that resulted from the changes in OBP. All of the calculations described were done by trial and error in an Excel spreadsheet. For each of the four cases (plus or minus 10 or 20 OBP points), the change in SLG needed to produce the same change in runs scored was divided by the OBP change to get an MVR, and then these four values were averaged to get an estimate for the specific baseline using the particular runs scored calculation method.

When Runs Created is used to calculate runs, the MVR estimates are 1.96 for the 2001 A's, 2.02 for the AL, and 2.01 for the NL in that year. When the Markov model is used, the estimates are A's: 2.10, AL 1.94, NL 1.95. Interestingly the relationships between the baselines are quite different between the two run calculation methods, although the quantitative differences are small. The values all suggest that 2 is a good simple estimate for the MVR.

One nice feature of the Markov model is that it enables calculation of expected scoring with nine different players, a real baseball lineup, rather than assuming all the players are the same, which is the case when the inputs are team or league data. Runs Created can't make an estimate based on the performance of the players taken individually rather than collectively. I decided to apply the Markov model to the lineup the A's used late in the season and for the playoffs. Because running the Markov model considering each player individually is somewhat complex and requires many trial and error runs, the only case I considered was when OBP was increased by 20 points. For that case, here are the MVRs by lineup position for the 2001 A's: (1) J. Damon, 2.80; (2) M. Tejada, 3.02; (3) Jason Giambi, 2.17; (4) J. Dye, 2.05; (5) E. Chavez, 1.91; (6) Jeremy Giambi, 2.13; (7) T. Long, 2.05; (8) R. Hernandez, 2.01; (9) F. Menechino, 2.47.

These values obviously depend on each player's position in the lineup and the abilities of the following hitters. Not surprisingly, the biggest relative payoff from increased OBP comes for batters in front of the power hitters, Damon and Tejada in the top two spots. By contrast, the lowest MVR is for Chavez in the number five slot who is followed by hitters without much power. Also contributing to his MVR are the high OBPs of Giambi (0.483) and Dye (0.373), so additional extra base hits by Chavez have a better chance of driving in runs. It is interesting to note that the lowest individual MVRs are around 2. That provides some evidence that perhaps 2 is on the low side. What needs to be done, and I hope to do it some day, is to study many more teams both in 2001 and recent years as well as in eras when scoring was much lower than it is now.

*Mark Pankin, mark@pankin.com* ♦

## Get Your Own Copy

If you're not a member of the Statistical Analysis Committee, you're probably reading a friend's copy of this issue of BTN, or perhaps you paid for a copy through the SABR office.

If that's the case, you might want to consider joining the Committee, which will get you an automatic subscription to BTN. There are no extra charges (besides the regular SABR membership fee) or obligations – just an interest in the statistical analysis of baseball.

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 (preferably with your snail mail address for our records) 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.

## Receive BTN by Internet Subscription

You can help save SABR some money, and me some time, by downloading your copy of *By the Numbers* from the web. BTN is posted to http://www.philbirnbaum.com in .PDF format, which will print to look exactly like the hard copy issue.

To read the .PDF document, you will need a copy of Adobe Acrobat Reader, which can be downloaded from www.adobe.com.

To get on the electronic subscription list, visit http://members.sabr.org, go to "My SABR," and join the Statistical Analysis Committee. You will then be notified via e-mail when the new issue is available for download.

If you don't have internet access, don't worry – you will always be entitled to receive BTN by mail, as usual.

# The Relative Value of OBA and SLG – Another View

Donald A. Coffin and Bruce W. Cowgill

*In a previous BTN, an article by Phil Birnbaum presented a method for determining how much a point of OBP is worth (in terms of run scoring) compared to a point of SLG. In this, one of two studies this issue on the topic, the authors revisit the question by examining empirical data on team offense.*

In the May, 2005 *BTN*, Phil Birnbaum examines the contention from *Moneyball*, made by Paul DePodesta, that one point of on-base percentage (OBP) is "worth" three points of slugging percentage (SLG). Birnbaum's analysis adjusts the 1987 AL data by raising SLG by one point, converting that into additional bases, and using Linear Weights to estimate the effect of that on runs scored. He compares this with two versions of increasing OBP (in one, the entire increase comes from walks, in the other, from hits), again calculating a (net) additional number of bases, and, applying Linear Weights, again calculated an effect on scoring. He concludes that one additional point of OBP is worth about 1.2 times as much as an additional point of SLG. A weakness of this approach is that it relies on an *a priori* knowledge of the "value" of an additional base.

## 1987-2004 MLB

An alternative approach is to use regression analysis. For all MLB teams, 1987-2004, we assembled their G, R, OBP, and SLG (acquired from the Sinins Sabermetric Encyclopedia). We calculated runs per game (RPG) (to adjust for the existing small variations in games played and the strike impacted years 1994 and 1995) and regressed RPG on OBP and SLG. The results are as follows:

**1987-2004 MLB**

| | | 95% Confidence Interval | |
|---|---|---|---|
| RPG = | t-statistic | Lower | Upper |
| - 5.354 | -31.37 | -5.689 | -5.019 |
| + 18.263 * OBP | +22.62 | 16.678 | 19.849 |
| + 9.611 * SLG | +25.20 | 8.862 | 10.360 |
| $r^2$ = 0.920 | F = 2889.873 | | |
| n = 506 | | | |
| | | Min | Max |
| Ratio of OBP to SLG: | 1.90 | 1.61 | 2.24 |

These results suggest that an increase in OBP of one point (e.g., from .340 to .341), holding SLG constant, increases RPG by 0.018263, while an increase in SLG by one point (e.g., from 0.400 to 0.401), holding OBP constant, increases RPG by 0.009611. The value, therefore, of OBP relative to SLG is 1.90. Using the 95% confidence intervals of the coefficients yields a ratio interval where the relative value of OBP ranges from 1.61 to 2.24.

From the high $r^2$ we can be confident in the model's predictive power. All of the coefficients are significant to the 99.9% level. Further, the largest difference between actual and predicted runs per game is 0.48. Out of the 506 team seasons analyzed, only one team season fell outside the three standard deviation statistical outlier "rule of thumb." This was the 2002 Phillies with 4.41 RPG, .339 OBP, .422 SLG. The model predicted the Phillies should have scored 4.89 RPG.

One issue is that the correlation between OBP and SLG is 0.80, which indicates that there is substantial multicollinearity between these two variables. This is not surprising, because batting average (BA) is a major component of both of these measures. One way to deal with this is to remove BA from one or the other of these two measures; this is easiest to do by calculating Isolated Power (ISO), which is SLG minus BA, and then regressing RPG on OBP and ISO. The results are shown below:

| 1987-2004 MLB | | | |
|---|---|---|---|
| | | 95% Confidence Interval | |
| RPG = | t-statistic | Lower | Upper |
| - 4.981 | -26.22 | -5.354 | -4.608 |
| + 24.713 * OBP | +36.44 | 23.380 | 26.045 |
| + 9.724 * ISO | +22.30 | 8.867 | 10.580 |
| $r^2$ = 0.909 | F = 2510.066 | | |
| n = 506 | | | |
| | | Min | Max |
| Ratio of OBP to ISO: | 2.54 | 2.21 | 2.94 |

Here, the coefficient of OBP is 2.5 times that of ISO, so the relative importance of a point of OBP appears larger. Using the 95% confidence intervals of the coefficients yields a ratio interval where the relative value of OBP ranges from 2.21 to 2.94. Oddly, however, the explanatory power of the regression declines.

## Is OBP more important in the AL or NL?

Performing the same analysis by league results in slightly different OBP relative values:

| 1987-2004 AL | | | |
|---|---|---|---|
| | | 95% Confidence Interval | |
| RPG = | t-statistic | Lower | Upper |
| - 5.406 | -22.72 | -5.875 | -4.938 |
| + 17.669 * OBP | +16.34 | 15.540 | 19.798 |
| + 10.271 * SLG | +19.64 | 9.241 | 11.301 |
| $r^2$ = 0.921 | F = 1449.731 | | |
| n = 252 | | | |
| | | Min | Max |
| Ratio of OBP to SLG: | 1.72 | 1.37 | 2.14 |

| 1987-2004 NL | | | |
|---|---|---|---|
| | | 95% Confidence Interval | |
| RPG = | t-statistic | Lower | Upper |
| - 5.004 | -19.79 | -5.502 | -4.506 |
| + 18.016 * OBP | 15.05 | 15.659 | 20.373 |
| + 8.887 * SLG | 16.42 | 7.821 | 9.953 |
| $r^2$ = 0.910 | F = 1273.947 | | |
| n = 254 | | | |
| | | Min | Max |
| Ratio of OBP to SLG: | 2.03 | 1.57 | 2.60 |

This means that a point of OBP is more valuable relative to a point of SLG in the NL than the AL. Or conversely, SLG is more valuable in the AL. During these years, the AL scored 4.83 RPG and the NL scored 4.51 RPG. With the higher scoring in the AL and the designated hitter, one could reasonably conclude that SLG is more valuable in the AL. To be clear, we are not saying that SLG is more valuable than OBP in the AL, just that a point of SLG is more valuable in the AL than a point of SLG in the NL.

## Does the relative value of OBP change from one season to the next?

Since the runs per game vary per year, we were curious how the value of OBP and SLG may also vary. From 1987 to 2004, runs per game ranged from a low of 4.1 in 1992 to a high of 5.1 in 2000. Applying our model to each year resulted in the following OBP relative values:

| Year | OBP / SLG | Year | OBP / SLG |
|---|---|---|---|
| 1987 | 2.46 | 1996 | 1.20 |
| 1988 | 2.46 | 1997 | 2.94 |
| 1989 | 0.76 | 1998 | 4.11 |
| 1990 | 0.37 | 1999 | 1.78 |
| 1991 | 1.74 | 2000 | 4.71 |
| 1992 | 1.90 | 2001 | 3.55 |
| 1993 | 1.06 | 2002 | 0.83 |
| 1994 | 1.26 | 2003 | 1.29 |
| 1995 | 2.51 | 2004 | 1.31 |
| | | 1987-2004 | 1.90 |

This is quite a range of values, from a low of 0.37 in 1990 to a high of 4.71 in 2000. A value under 1.00 would indicate that a point of SLG is worth more than a point of OBP. We were surprised at such a result as we expected the ratios to be much closer during this time period. That is, we could understand different results if comparing today's game to the 1960s or the deadball era, but not within the same era.

While the sample sizes for each season are small (26-30 teams per season), the $r^2$ of each model is still fairly high, ranging from .807 to .948. The coefficients are also significant.

In fact, such variation from season to season is not particularly surprising or even unusual. The sample sizes are small, making the coefficient estimates less precise. We performed a Chow Test on the relationship between the individual season regressions and the pooled regression. The null hypothesis of the Chow Test is that the individual regressions are not statistically distinguishable from the pooled regression. We find that we cannot reject the null hypothesis, which suggests that the pooled regression is the "best" explanatory model. Essentially, using individual season data is like taking a sub-sample of the entire data set and running the regression on it. (However, the seasons are not random samples from the entire data set.) These smaller samples will lead to coefficient estimates that differ from the pooled data set, but the Chow Test suggests that the different coefficient estimates are not systematic.

There does not seem to be any type of pattern in the results. For instance, recent seasons vary as much as the late 1980s or the 1990s. There is little correlation between RPG and OBP's relative value (r = 0.38), meaning that a point of OBP is more valuable than SLG in some high RPG seasons as well as some low RPG seasons. Therefore, we cannot reasonably conclude that OBP or SLG is more valuable in particular run environments.

We even controlled for multicollinearity by examining OBP and ISO. The results still show large differences season to season.

| Year | OBP / ISO | Year | OBP / ISO |
|---|---|---|---|
| 1987 | 2.73 | 1996 | 2.06 |
| 1988 | 3.07 | 1997 | 3.23 |
| 1989 | 1.40 | 1998 | 4.65 |
| 1990 | 0.74 | 1999 | 2.13 |
| 1991 | 2.11 | 2000 | 8.73 |
| 1992 | 2.36 | 2001 | 4.26 |
| 1993 | 1.25 | 2002 | 1.91 |
| 1994 | 1.68 | 2003 | 2.19 |
| 1995 | 3.32 | 2004 | 2.17 |
| | | 1987-2004 | 2.54 |

## Summary

Birnbaum's OBP relative value of 1.20 is less than the 2.46 value we found for 1987. It is even lower than our 18-season model's 1.90 figure. Birnbaum's initial analysis was based on linear weights figures that might undervalue certain events. Birnbaum recently adjusted his calculation by using the season specific linear weights and the value of OBP to SLG increased to 1.43[1]. To be fair, Birnbaum's figure falls within our model's confidence interval. In 1987, our model's ratio of 2.46 falls within a ratio interval ranging from 1.08 to 5.84 (quite large due to a sample size of 26, $r^2$=.816).

Mark Pankin's analysis in this issue results in a ratio of about 2.0. Ray Sauer, in a paper with Jahn Hakes,[2] also puts the value around 2.0. And Paul DePodesta thinks it is 3.0. There may be reason to believe that the real effect of differences between teams in SLG actually has a larger effect on scoring, because there is greater variation (between teams) in SLG than there is in OBP.

We expect different calculations and assumptions to result in different values. Yet, even though the results vary depending on method and dataset, each analysis seems to arrive at similar conclusions: One point of OBP is more valuable than one point of SLG.

## Cautionary Note

One thing of concern in this type of analysis is that it may imply (or seem to imply) that getting on base is more important than slugging. (We are speaking of the acts themselves, not the statistics.) While it is true that a point of OBP is more valuable than a point of SLG (regardless of whether it is 1.5 times or 2 times as important), the fact that each stat is scaled differently is the reason that this occurs and not the relative value of the underlying events. That is, the actual value of getting on base as compared to advancing runners is the opposite of what this type of analysis may lead some to think. Since OBP is based on a 0 to 1 scale and SLG is based on a 0 to 4 scale, the results may suggest (to the uninformed) that getting on base is more important than the type of hit. If we scale SLG on a 0 to 1 scale (simply by dividing it by 4), we get an entirely different picture. Using the 1987-2004 data, we saw that OBP is 1.9 times as valuable as SLG. But, if we "normalize" SLG to the same scale as OBP, the relationship is almost exactly reversed -- the "new" SLG becomes 2.1 times as valuable as OBP. Of course, linear weights and other systems clearly show that extra bases are highly valued, so we are not stating that there was any doubt in the value system. It is just that the *Moneyball* statement that OBP is more important than SLG may cause some to misinterpret what that really means. *Moneyball's* underlying message is not that getting on base is *the* most important thing or more important than slugging; rather, it asserts that getting on base has been undervalued.

*Don Coffin, dcoffin@iun.edu*
*Bruce Cowgill, brucec23@comcast.net* ♦

[1] See *Phil Birnbaum Responds*, page 14.
[2] http://hubcap.clemson.edu/~sauerr/working/moneyball-v2.pdf