

Blog Export: Beaneball, <http://beaneball.org/>

Friday, September 17, 2004

MVP Metrics

A few days ago, Studes at Hardball Times made the case that Win Shares should be the definitive statistic for deciding who the MVP is. He makes a compelling argument, but what it brings up for me is the necessity of something better, something that uses the outcome of each individual at-bat for every player to determine their full worth.

The thing is, that stat has been created. Jim Albert, author of *Curve Ball: Baseball, Statistics, and the Role of Chance in the Game*, wrote this paper describing how he took play-by-play data from the 1987 National League to determine the value of each hitter's contributions to his team.

The idea is that every base-out state of a game has an expected run value attached to it. A player's at-bat will change that base-out state, and thus the expected runs for that inning. The player, then, is given credit (or debit) for the effect of his at-bat. For example, if Barry Bonds comes up with one out and a runner on second, he can expect his team to score 0.68 runs. When he, as usual, walks, the state is now one out and runners on first and second. That position results in 0.86 expected runs for the Giants. Thus, Barry has 0.18 runs added to his total.

The point, as Studes makes in his argument for Win Shares, is that, while the player's actual value may not be predictive (because the base-out situations they face may vary wildly from year to year), their performance in different states does affect their value. Just because, for example, Timo Perez isn't going to hit as well with runners in scoring position next year as he has this time around doesn't make his accomplishment any less valuable.

Surely, the sabermetric community knows about Albert's paper. It is dated September 9th, 2001. Understandably, we were preoccupied with other things very soon after Albert put his paper up, but in the years since, why hasn't a site cropped up using his basic formula to calculate the actual values of players? (This assumes, of course, that there hasn't been just this effort made and I don't know about it.)

My guess is the lack of available data. Retrosheet, of course, is a marvelous solution for old data. But how does one get a hold of daily play-by-play transcripts? Baseball Info Solutions is one place, I suppose, but you pay for the data.

Perhaps I'm missing something. Perhaps there is some place where enterprising sabermetricians get their hands on all the data they need for free. Did this clearinghouse exist, you could bet on finding the coolest value metric around updated daily on this site.

Posted by jason in Baseball at 19:57

hi jason!

great blog!

i agree that arnold's way sounds more sensible, except it is offense only. i object to win shares because i object to basing it on imaginary wins and losses (the pythagorean formula.)

i think that too many people just worship bill james and agree with anything he sez.

lisa

Anonymous on Sep 17 2004, 22:14