By the Numbers

Comment

Academic Research: Organizational Decisions
Charlie Pavitt

The author reviews three recent papers that focus on various aspects of player evaluation and promotion.

Tony Caporale and Trevor C. Collier (2013), <u>Scouts versus stats: The impact of *Moneyball* on the major league baseball draft</u>, Applied Economics, Vol. 45, 1983-1990

Gabriel Chandler and Guy Stevens (2012). <u>An exploratory study of minor league baseball statistics,</u> Journal of Quantitative Analysis in Sports, Vol. 8 Issue 4, Article 4

Opedace, Robert, and Janet Kiholmsmith (2013). <u>Loss aversion and managerial decisions: Evidence from major league baseball,</u> Economic Inquiry, Vol. 51 No. 2, pp. 1475-1488

In my opinion, there hasn't been any truly critical statistical baseball research from the academic side over the past couple of years. There has been minor work that may be of some interest. I've chosen to summarize three articles that focus in one way or another on organizational decision-making

During the period *Moneyball* covers, Billy Beane and his compatriots concentrated on drafting from the college ranks rather than high school, having concluded that the former were undervalued relative to the latter in the baseball labor market. Some past academic research has been consistent with that. Here, Caporale and Collier revisit the topic by examining whether first round draft choices out of high school or college had the more successful careers.

The authors used data only from 1995 through 1999 in order to pre-date the publication of the book and any influence it might have had in other teams' drafting strategies, providing them with too small a sample size in my view. Their estimate for career Wins Above Replacement was in line with college players performing better, but was not statistically significant. They corrected for the fact that players never making the majors have zero WAR, but (unless I missed it) not for the odds of the draftees making it in the first place.

They also noted a significant tendency for players drafted earlier in the first round to be more likely to make the majors, estimating the odds decreasing by two percent for each lower draft position. Finally, looking more globally at the first three rounds from 1965 to 2010, there was a trend for a smaller percentage of draftees to be high schoolers over time, although no evidence that the publication of *Moneyball* instigated or hastened that tendency.

Chandler and Stevens attempted to determine the factors that organizations appear to use in promoting young position players through the different levels of the minors and eventually to a significant career in the majors, defined as at least 320 games, the equivalent to two full seasons. Their original sample included all 1019 draftees between 1999 and 2002 who appeared in at least 15 minor league games, along with a secondary sample of draftees from 1995 to 1998. Basically, their aim was to see how well performance as measured by different standard indices at each of six levels (Rookie, low A, A, high A, AA, AAA) predicted an eventual major league career.

First, and not surprisingly, overall prediction improved markedly at higher levels over lower levels. What is most interesting are

the differing impacts of various indices across levels. The standard rate indices of BA, OBA, SA, and OPS were predictive at AA and AAA, but not lower than that. Walks per at-bat and RBI were not predictive at any level.

At low levels one factor that does matter is draft position --organizations are willing to give the benefit of the doubt to early picks. Interestingly, and plausibly, once one thinks about it, strikeouts per at bat and per walk are diagnostic for early draft picks; apparently, they are the best indicator of whether a player is overmatched.

Opedace and Kiholmsmith attempt to use baseball to check whether managers are more likely to terminate poor performers if those performers were originally hired by someone than themselves. Relying on data from 1976 through 2005, they determined that poorly performing players, defined as those in the bottom one-fourth in both at bats and slugging average, are particularly likely to leave a team when that team's current general manager is in his first or second year of tenure and that player was acquired by the immediately previous general manager.

The authors believe that when a GM adds a player to the roster who plays badly, he is likely to retain that player rather than admit the error; the next GM has no such qualms about jettisoning the player.

Unfortunately, Opedace and Kiholmsmith's data set had two compromising weaknesses: it did not distinguish players who chose to leave a team from players whom the team no longer wanted, and it had no qualifications for contract conditions other than the existence of a no-trade clause. The authors recognized the second problem, and responded that their findings were substantially the same when limited to players' first six seasons, when free agency was not a possibility.

Charlie Pavitt, chazzq@UDel.Edu ♦

The 1971 Baltimore Orioles

Tom Hanrahan

What was behind the success of the 1971 Baltimore Orioles? Was it great pitching, as conventional wisdom would have it, or something else? The author investigates.

The 1971 Baltimore Orioles are the answer to a classic trivia question: "Name the last team to have four 20-game winning pitchers." This article examines the team; what was behind their success?

The O's from 1964 to 1983 were a dynasty that lasted a generation. Over those 20 years, they averaged over 93 wins per full season. They won six pennants and three World Series. Only once did they finish below .500, and only once did they not score more runs than they allowed. Many times they won the division or pennant (before divisions existed) by huge margins.

Within this long string of success, the team in 1969-1971 is often thought of as one of the best mini-dynasties of all time. Three division titles. A perfect 9-0 in the ALCS. Three straight pennants. Three straight years of over 100 wins. Every year, arguably the best offense and the best defense in the American League . Only the losses in two World Series blemish their mark -- one loss to the Amazing '69 Mets, who played out of their socks (and were helped by the shoe-polish HBP); and the other to the '71 Pirates.

The 1971 O's, at 101-57, did not have quite the gaudy W-L record of the previous two years (108 and 109 wins), but they were still very dominant. The Orioles of 1969-71 might be remembered as the greatest brief dynasty of all, had they won the '71 World Series. Incredibly, one of their greatest strengths abandoned them. The O's, completely out of character, made 9 errors in the first 5 games of the Series; the majority of them by Mark Belanger and Brooks Robinson (!).

What We Think We Know

Well, they had four 20-game winners: Jim Palmer, Mike Cuellar, Dave McNally, and Pat Dobson. The team had the league's best ERA. The O's are remembered as having a fine defense, and overall a solid lineup with two Hall of Fame hitters named "Robinson". However, no batter had a particularly great year. The only Oriole hitter to lead the league in a category was Buford's 99 runs scored. No batter drove in 100 runs. No one finished in the top 10 in total bases. As a team, their hitters did not lead the league in the traditional categories of highest batting average or most home runs. So, on the surface, it seems to be a team of four great pitchers and a fine supporting cast. So it seems.

A View From the Top

What do we see if we compare the team's stats with the rest of the American League? So glad that you asked.

	runs scored	runs allowed
Orioles	742	530
League average	623	623
Difference	119	93
Ratio	1.191	1.175

Okay, so the O's were excellent on both sides. If anything, their ability to score runs was stronger than their ability to prevent them. Their park was basically neutral for run scoring, so there are no distorting effects there. On the surface, this would seem to imply that we ought to give approximately equal credit to the pitching and the hitting; and of course the pitching consists mostly of 4 men, while the batting is distributed among many. That, I will argue, would be wrong. It *is* true that there should be almost as much credit given to the run-preventing exploits of the '71 O's as the run-producing efforts. However, run prevention is made up of two items; pitching, and defense. So we will need to assess how much of the run-prevention credit should go to the gloves, as opposed to the arms.

Position Players

The O's basically played 1970 with 10 men. They platooned at catcher, with LH hitting Elrod Hendricks getting a bit more playing time than Andy Etchebarren (who still started 61 games behind the plate). Their infield was set; almost every day, it was Davey Johnson at 2B, with Belanger and Robby on the left side. Boog Powell played first base, although at times Frank Robinson moved from RF to play 1B . Don Buford played LF, Paul Blair CF, and Merv Rettenmund played everywhere in the OF; they basically had 5 men alternating between 4 positions, with each of them coming to bat over 500 times. This was a team that could have been even better with a DH!

Aside from those ten players, no one else got on to the field very much; in fact, if you order the team by most times at bat, the next three men after these ten would be pitchers. The team had young stars Bobby Grich and Don Baylor sitting on the bench, waiting for the future to get their chances to play. As far as league leaders go, aside from Buford mentioned above, Rettenmund finished third in batting average, and Frank Robinson was among the leaders in HR, slugging, and RBI. Brooks Robinson had his typical year; three of the 24 voters for the MVP award chose the Human Vacuum Cleaner. The key to the O's offense was that nobody came close to having a *bad* year, or even a mediocre one; like the fictional Lake Wobegon, they were all above average. The worst batting average among the eight fulltime hitters was Powell's .256, but he drove in 92 runs. Belanger was the only man without power. Every other hitter slugged over .390, in a league where the mean SLG was .364. Even Belanger got on base at an excellent .365 rate, which was much better than his typical season.

		AB	avg	HR	R	RBI	BB	OPS
Elrod Hendricks	C	316	.256	9	33	42	39	.720
Andy Etchebarren	C	243	.270	9	21	29	16	.749
Boog Powell	1B	418	.256	22	59	92	82	.839
Davey Johnson	2B	510	.282	18	67	72	51	.794
Brooks Robinson	3B	589	.272	20	67	92	63	.754
Mark Belanger	SS	500	.266	0	67	35	73	.685
Frank Robinson	RF-1B	455	.281	28	82	99	72	.894
Don Buford	LF	449	.290	19	99	54	89	.890
Paul Blair	CF	516	.262	10	75	44	32	.703
Merv Rettenmund	OF	589	.318	11	81	75	87	.870

Pitchers

The O's had a 4-man rotation. And…. the O's had a 4-man rotation. Oh, and when McNally was injured, Grant Jackson briefly turned from reliever into starter. Aside from that, the O's, well, they had a 4-man rotation. The big four tossed seventy complete games. No pitcher on the team threw even 45 innings in relief. In the post-season, the top 3 starters tossed 70 innings, while the rest of the team combined was used for 18.

While the team is famous for its 20-game winning rotation, none of the pitchers had a singularly great season; however, they were consistently good. Their records were as follows:

	W	L	ERA
Jim Palmer	20	9	2.68
Mike Cuellar	20	9	3.08
Dave McNally	21	5	2.89
Pat Dobson	20	8	2.90

The rest of the pitching staff had a W-L record of 20-26 with a 3.31 ERA. Stated a different way, everyone else combined got credit for the other fifth of the Orioles total of 100+ victories. The Big Four gained all of their "W"s as starters. Only Dobson pitched at all in relief, with one appearance (a save).

All four starters received some MVP voting consideration. McNally finished 11th in the voting, Dobson 17th, Palmer 23rd, and Cuellar 27th. McNally finished tied for fourth in Wins, while the others were tied for sixth. Palmer was third in the league in ERA; McNally and Dobson were seventh and eighth.

Who were these four men? Did this glut of quality come out of nowhere? Or was it a case of stars aligning for a talented group?

Jim Palmer's career was so strong that, if anything, 1971 was a bit of an off year for him. The man who would win three Cy Young awards and lead the league in many categories in his career, just did in '71 what he would often do (7 times) in his career; won 20 games.

Dave McNally won 184 games in his career, losing 119. This is impressive, but his winning percentage largely reflects the teams for which he played. He was basically a somewhat-above-average pitcher who was done by age 35.

Mike Cuellar was a fine pitcher, toiling away for the expansion Houston Astros, and not winning many games there. He was traded to the Orioles in the fall of 1968, and immediately proceeded to win 20 in each of the next three years. In the previous season, 1970, Cuellar had finished 24-8, despite giving up more home runs and more earned runs than any other pitcher in the American League. Cuellar did not get his first MLB win until he was 27, but he lasted until 40, and finished with a similar record to McNally: 185-130.

Pat Dobson lost more games than he won in his career as a whole. But in 1971, he pitched well, and the Orioles consistently scored runs for him. In his 37 starts, only 5 times did Baltimore score fewer than three runs.

Where does that leave us? Eight fine hitters (plus the catching platoon) and four ace arms. How to go about apportioning credit?

Credit for Run Prevention

Let's compare the Orioles team offense and team defense in several categories:

	AB	H	BB	1B	2B	3B	HR	R	K	outs	IPouts	BIP	BABIP
offense	5303	1382	672	992	207	25	158	742	844	3921	3077	4301	.285
defense	5268	1257	416	922	189	21	125	530	793	4011	3218	4350	.260
difference	35	125	256	70	18	4	33	212	51	−90	−141	−49	.025
league avg	5387	1330	540	975	202	29	124	623	868	4057	3189	4395	.274

(Notes:
1. "outs" is simply at bats minus hits (I did not add in SF, SH, CS, GIDP, outs on bases, etc.
2. "inplay outs" is all outs minus strikeouts.
3. BIP is balls in play; all official bats not resulting in strike outs or home runs.
4. BABIP is batting average on balls in play, calculated as (hits minus HR) / (balls in play).)

The Orioles outscored their opponents by 212 runs. This can be attributed to a few factors:

- They drew 256 more walks than their pitchers allowed. That, my friends, is an impressive and valuable team achievement.
- They hit 33 more home runs than their pitchers allowed.
- They hit 92 more { singles+doubles+triples } than they allowed; this in spite of the fact that their batters struck out 51 more times than their opponents.

Looking at the Orioles' run prevention, we see that

- Their pitchers allowed 124 fewer walks than average
- But they struck out 75 fewer men than average
- They allowed almost exactly an average number of home runs

So they were very stingy with the walks. Whether you attribute this to Earl Weaver's insistence on throwing strikes, the moundsmen's confidence in the gloves behind them, or simply a good set of control artists, we should give the hurlers credit for not giving out free passes. Other than that, though, the team's pitchers were mediocre. They allowed fewer non-homerun hits than most teams; but how much to credit that to the pitchers, and how much to the defense? We've had oodles of metrics in the past generation that have attempted to do this, and I'll refer to those in a bit. But it should be obvious to anyone who watched this team that it was composed of an unusually gifted set of defenders. Four men won gold gloves in 1971; Blair, Belanger, Brooks, and Johnson. The first three of those are considered by many to be

among the best ever gloves at their positions. The team made few errors (second least in the league) and allowed few unearned runs. Many would argue, even without referring to advanced statistics, that much of the credit to Oriole run prevention, which 93 runs better than league average, should go to the men flashing the leather. Are we able to quantify this? Standing on the shoulders of giants, we can try.

Credit Apportioned to Players

I will use here two common, well-accepted meta-systems that attempt to take the entirety of players' accomplishments and express them in terms of a number of wins.

The first is Win Shares (WS), which begins with a top-down approach; given how many games a team won, it divides the credit among the individuals. The second is Wins Above Replacement (WAR), as calculated and viewable on baseball-reference.com, which analyzes players' accomplishments individually.

One major difference between the two metrics is that if a team "over-performs" its individual parts (that is, if it wins more games than one would expect a team to have won, given its players contributions, possibly through clutch play), this credit is given to the players in WS, but is attributed to "luck" and therefore lost using WAR. The 1971 O's were not significantly affected by this; their record of 101-57 is about what would be expected of a team with their underlying statistics of hits, walks, etc.

A second difference between the two systems is that WS gives more credit to average performance than WAR. Another way to phrase it is that the "replacement level" is set much lower using WS. One cannot achieve negative win shares (that is, a number below zero), but one can, if performing very poorly, wind up with negative WAR. Therefore, the wins totals for individuals are higher. This does not have any effect when comparing the Rank of player values; it merely shifts the values in one system's baseline all in one direction.

Lastly, because WS as invented by Bill James are denominated into "thirds of wins", I divided all WS by 3 to create an actual wins total for comparison to WAR.

The O's team had 101 WS (by definition; they won 101 games) and 53.4 WAR. The main 14 players were worth 91 of the WS, and 52.6 of the WAR; very little credit is due any of the others (both systems would see Eddie Watt, who led the team with 11 saves, being the 15th most valuable Oriole). The big four pitchers were worth 26% and 24%, in each system respectively, of their teams' totals.

The following table compares those players' values as determined by WS and WAR. The players in the table are grouped by the four pitchers, followed by the ten hitters, ordered within their group by the average of WS and WAR they achieved.

	Win Shares (divided by 3)	WS rank on the team	Wins Above Replacement	WAR rank on the team	average of WS and WAR
Palmer	7.3	6	4.1	6	5.7
Dobson	6.3	9	3.1	9	4.7
McNally	6.3	9	3.0	10	4.7
Cuellar	6.0	11	2.8	12	4.4
Rettenmund	9.0	1	5.9	2	7.5
Buford	8.7	2	5.1	3	6.9
B Robinson	7.7	4	6.0	1	6.8
Johnson	7.7	4	4.4	5	6.0
Belanger	7.0	7	4.6	4	5.8
F Robinson	7.7	4	3.3	8	5.5
Powell	6.3	9	3.6	7	5.0
Blair	5.0	12	2.9	11	4.0
Etchebarren	3.3	13	1.8	14	2.6
Hendricks	2.7	14	2.0	13	2.3

Conclusion

So there you have it. The 1971 Orioles were led by three Hall of Famers, with four gold gloves and four twenty-game winners. But, in the end, their greatest contributor was Merv Rettenmund -- a man who was having his only year in which he played close to full time. Rettenmund's .422 on-base percentage went mostly unnoticed, as he finished sixth among his teammates in MVP consideration.

The second-most valuable Oriole was Don Buford, who received even less credit than Rettenmund despite leading the league in runs scored. Buford, age 34, would completely fall apart the next season, hitting .206 and subsequently retiring.

The big four pitchers were solid and durable, but none of them were even among the five most valuable players on the team. Those arms each won 20 games by the skin of their teeth; any different distribution of runs scored by their offense, or lesser plays by the defense, could easily have resulted in one of them winning 24, but a team of possibly three, two, one, or even zero 20-game winners. By a fluke of events, the trivia question is cemented in history. However, the reality of the team's success is much different than the factoid suggests.

The reality is the O's were a *team*. They had 12 full time players (eight bats and four arms) plus a fine catching platoon, and no weaknesses whatsoever. Zero. How many teams have ever had their primary 14 players be above average? I don't know, but the odds of a fair coin coming up heads fourteen times in a row are 1 in 16,384, and there have not been nearly that many teams in the history of MLB.

The 1971 Orioles are known by many as the answer to a trivia question, but not known for what they should be: one of the most complete *teams* of all time.

Tom Hanrahan, han60man@aol.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, 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 .

"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.